

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

Regulatory DNA sequences of the gene for the human catalytic telomerase subunit, and their diagnostic and therapeutic use

Structure and function of the chromosome ends

5

The genetic material of eukaryotic cells is distributed on linear chromosomes. The ends of hereditary units are termed telomeres, derived from the Greek words *telos* (end) and *meros* (part, segment). Most telomeres consist of repeats of short sequences which are mainly composed of thymine and guanine (Zakian, 1995). In all the vertebrates which have so far been investigated, the telomeres consist of the sequence TTAGGG (Meyne *et al.*, 1989).

10

15

The telomeres have a variety of important functions. They prevent the fusion of chromosomes (McClintock, 1941) and thus the formation of dicentric hereditary units. Such chromosomes having two centromeres can lead to the development of cancer due to loss of heterozygosis or duplication, or loss of genes.

20

In addition, telomeres serve the purpose of distinguishing intact hereditary units from damaged hereditary units. Thus, yeast cells ceased their cell division when they contained a chromosome without a telomere (Sandell and Zakian, 1993).

25

Telomeres fulfil another important task in association with the replication of eukaryotic cell DNA. In contrast to the circular genomes of prokaryotes, the linear chromosomes of eukaryotes cannot be completely replicated by the DNA polymerase complex. RNA primers are required to initiate DNA replication. After elimination of the RNA primers, extension of the Okazaki fragments and subsequent ligation, the newly synthesized DNA strand lacks the 5' end since the RNA primer cannot be replaced by DNA at that point. Without special protective mechanisms, the chromosomes would therefore shrink with each cell division ("end-replication problem"; Harley *et al.*, 1990). The non-coding telomere sequences presumably constitute a buffer zone for preventing the loss of genes (Sandell and Zakian, 1993).

30

001260 "STC" 830

In addition to this, telomeres also play an import role in regulating cell ageing (Olovnikov, 1973). Human somatic cells exhibit a limited capacity for replication in culture; after a certain period of time, they become senescent. In this state, the cells no longer divide even after having been stimulated with growth factors; however, they do not die and remain metabolically active (Goldstein, 1990). Various observations support the hypothesis that a cell determines how many more times it can divide on the basis of the length of its telomeres (Allsopp *et al.*, 1992).

In summary, the telomeres consequently possess key functions in the ageing of cells, and in stabilizing the genetic material and preventing cancer.

The enzyme telomerase synthesizes the telomeres

As described above, organisms which possess linear chromosomes can only replicate their genome incompletely in the absence of a special protective mechanism. Most eukaryotes use a special enzyme, i.e. telomerase, for regenerating the telomere sequences. Telomerase is expressed constitutively in the single-cell organisms which have so far been investigated. On the other hand, telomerase activity has only been measured in humans in germ cells and tumour cells, whereas neighbouring somatic tissue did not contain any telomerase (Kim *et al.*, 1994).

Telomerase can also be designated functionally as terminal telomere transferase, which is located in the cell nucleus as a multiprotein complex. While the RNA moiety of human telomerase has been known for a relatively long period of time (Feng *et al.*, 1995), the catalytic subunit of this enzyme group was recently identified in a variety of organisms (Lingner *et al.*, 1997; cf. our application PCT EP/98/03468 which is likewise pending). These catalytic subunits of telomerase are strikingly homologous both among themselves and in relation to all previously known reverse transcriptases.

WO 98/14592 also describes nucleic acid and amino acid sequences of the catalytic telomerase subunit.

Activation of telomerase in human tumours

It was originally only possible to demonstrate telomerase activity in humans in germ
line cells and not in normal somatic cells (Hastie *et al.*, 1990; Kim *et al.*, 1994).
Following the development of a more sensitive detection method (Kim *et al.*, 1994),
a low telomerase activity was also detected in hematopoietic cells (Broccoli *et al.*,
1995; Counter *et al.*, 1995; Hiyama *et al.*, 1995). It is true, however, that these cells
nevertheless exhibited a reduction in the telomeres (Vaziri *et al.*, 1994; Counter *et al.*,
1995). It has still not been resolved whether the quantity of enzyme in these cells
is not sufficient for compensating the telomere loss or whether the telomerase activity
which is measured stems from a subpopulation, e.g. incompletely differentiated
CD34⁺38⁺ precursor cells (Hiyama *et al.*, 1995). In order to resolve this, it would be
necessary to detect telomerase activity in a single cell.

Interestingly, however, significant telomerase activity was detected in a large number
of the tumour tissues which had thus far been tested (1734/2031, 85%; Shay, 1997),
whereas no activity was found in normal somatic tissue (1/196, <1%, Shay, 1997). In
addition various investigations have shown that the telomeres still shrank in
senescent cells which were transformed with viral oncoproteins and it was only
possible to detect telomerase in the subpopulation which survived the growth crisis
(Counter *et al.*, 1992). The telomeres were also stable in these immortalized cells.
(Counter *et al.*, 1992). Similar findings from investigations in mice (Blasco *et al.*,
1996) support the assumption that reactivation of the telomerase is a late event in
tumorigenesis.

Based on these results, a "telomerase hypothesis" was developed which links the loss
of telomere sequences and cell ageing with telomerase activity and the development
of cancer. In long-lived species such as humans, the shrinking of the telomeres can be
regarded as being a mechanism for suppressing tumours. Differentiated cells which
do not contain any telomerase cease their cell division at a particular telomere length.
If such a cell mutates, it can only form a tumour if the cell can extend its telomeres.

Otherwise, the cell would continue to lose telomere sequences until its chromosomes became unstable and it was finally destroyed. Telomerase reactivation is presumably the main mechanism used by tumour cells to stabilize their telomeres.

5 It follows from these observations and considerations that it should be possible to treat tumours by inhibiting the telomerase. Conventional cancer therapies using cytostatic agents or short-wave radiation damage all the dividing cells in the body in addition to the tumour cells. However, since only germ line cells, apart from tumour cells, contain significant telomerase activity, telomerase inhibitors would attack the
10 tumour cells more specifically and consequently elicit fewer undesirable side effects. Telomerase activity has been detected in all the tumour tissues which have so far been tested, which means that these therapeutic agents could be employed against all types of cancer. The effect of telomerase inhibitors would then set in when the telomeres of the cells had shortened to such an extent that the genome became
15 unstable. Since tumour cells usually possess telomeres which are shorter than those of normal somatic cells, cancer cells would be the first to be eliminated by the telomerase inhibitors. By contrast, cells possessing long telomeres, such as the germ cells, would only be damaged at a much later date. Telomerase inhibitors consequently represent a potential way forward in the treatment of cancer.

20

It becomes possible to obtain unambiguous answers to the question of the nature and points of attack of physiological telomerase inhibitors once the manner in which expression of the telomerase gene is regulated has also been identified.

25 Regulation of gene expression in eukaryotes

There are a large number of points in eukaryotic gene expression, i.e. the cellular flow of information from the DNA to the protein by way of the RNA, at which regulatory mechanisms can exert an effect. Examples of individual control steps are
30 gene amplification, the recombination of gene loci, chromatin structure, DNA methylation, transcription, post-transcriptional modifications of mRNA, mRNA transport, translation and post-translational modifications of proteins. Studies which

have been carried out to date indicate that control at the level of transcription initiation is of the greatest importance (Latchman, 1991).

5 A region which is responsible for regulating transcription, and which is designated the promoter region, is located directly upstream of the transcription start of a gene which is transcribed by RNA polymerase II. Comparison of the nucleotide sequences of promoter regions from a large number of known genes shows that particular sequence motifs occur regularly in this region. These elements include, inter alia, the TATA box, the CCAAT box and the GC box, which elements are recognized by
10 specific proteins. The TATA box, which is located about 30 nucleotides upstream of the transcription start, is, for example, recognized by the TFIID subunit TBP ("TATA box-binding protein"), whereas particular GC-rich sequence segments are specifically bound by the transcription factor Sp1 ("specificity protein 1").

15 The promoter can be functionally subdivided into a regulatory segment and a constitutive segment (Latchman, 1991). The constitutive control region comprises the so-called core promoter which enables transcription to be initiated correctly. This promoter contains the sequence elements which are described as UPE's (upstream promoter elements) which are necessary for efficient transcription. The regulatory
20 control segments, which can be interlaced with the UPE's, possess sequence elements which can be involved in the signal-dependent regulation of transcription by hormones, growth factors, etc. They impart tissue-specific or cell-specific promoter properties.

25 DNA segments which are able to exert an influence on gene expression over relatively large distances are a characteristic feature of eukaryotic genes. These elements can be located upstream or downstream of a transcription unit, or within the unit, and can perform their function independently of their orientation. These sequence segments may reinforce (enhancers) or attenuate (silencers) promoter
30 activity. In a similar way to the promoter regions, enhancers and silencers also accommodate several binding sites for transcription factors.

The invention relates to the DNA sequences from the 5'-flanking region of the gene for the catalytically active human telomerase subunit and intron sequences for this gene.

5 The invention particularly relates to the 5'-flanking regulatory DNA sequence which contains the promoter DNA sequence for the gene for the human catalytic telomerase subunit, as depicted in Fig. 10 (SEQ ID NO 3).

10 The invention furthermore relates to part regions of the 5'-flanking regulatory DNA sequence, as depicted in Fig. 4 (SEQ ID NO 1), which has a regulatory effect.

15 Intron sequences for the gene for the human catalytic telomerase subunit, in particular those sequences which have a regulatory effect, are also part of the subject-matter of the present invention. The intron sequences according to the invention are described in detail in the context of Example 5 (cf. SEQ ID NO 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20).

20 The invention furthermore relates to a recombinant construct which comprises the DNA sequences according to the invention, in particular the 5'-flanking DNA sequence of the gene for the human catalytic telomerase subunit, or part regions thereof.

25 Preference is given to recombinant constructs which, in addition to the DNA sequences according to the invention, in particular the 5'-flanking DNA sequence of the gene for the human catalytic telomerase subunit, or part regions thereof, also contain one or more additional DNA sequences which encode polypeptides or proteins.

30 According to a particularly preferred embodiment, these additional DNA sequences encode antineoplastic proteins.

5

10

15

20

25

30

herpes simplex virus thymidine kinase, varicella zoster virus thymidine kinase, bacterial nitroreductase, bacterial β -glucuronidase, plant β -glucuronidase from *Secale cereale*, human glucuronidase, human carboxypeptidase, bacterial carboxypeptidase, bacterial β -lactamase, bacterial cytosine deaminidase, human catalase and/or phosphatase, human alkaline phosphatase, type 5 acid phosphatase, human

lysooxidase, human acid D-aminooxidase, human glutathione peroxidase, human eosinophil peroxidase and human thyroid peroxidase.

5 The abovementioned recombinant constructs can also contain DNA sequences which encode factor VIII or factor IX, or part fragments thereof. These DNA sequences also include other blood clotting factors.

The abovementioned recombinant constructs can also contain DNA sequences which encode a reporter protein. Examples of these reporter proteins are:

10 Chloramphenicol acetyl transferase (CAT), glow-worm luciferase (LUC), β -galactosidase (β -Gal), secreted alkaline phosphatase (SEAP), human growth hormone (hGH), β -glucuronidase (GUS), green-fluorescing protein (GFP), and all the variants derived therefrom, aquarin and obelin.

15 Recombinant constructs according to the invention can also contain DNA which encodes the human catalytic telomerase subunit and its variants and fragments in the antisense orientation. Where appropriate, these constructs can also contain other protein subunits of the human telomerase and the telomerase RNA component in the
20 antisense orientation.

The recombinant constructs can, in addition to the DNA which encodes the human catalytic telomerase subunit, and its variants and fragments, also contain other protein subunits of the human telomerase and the telomerase RNA component.

25 The invention furthermore relates to a vector which contains the abovementioned DNA sequences according to the invention, in particular the 5'-flanking DNA sequences and also one or more of the other DNA sequences mentioned above.

30 The preferred vector for these constructs is a virus, for example a retrovirus, an adenovirus, an adeno-associated virus, a herpes simplex virus, a vaccina virus, a lentiviral virus, a Sindbis virus and a Semliki forest virus.

Preference is also given to using plasmids as vectors.

5 The invention furthermore relates to pharmaceutical preparations which comprise recombinant constructs or vectors according to the invention; for example a preparation in a colloidal dispersion system.

Examples of suitable colloidal dispersion systems are liposomes or polylysine ligands.

10

The preparations of the constructs or vectors according to the invention in colloidal dispersion systems can be supplemented with a ligand which binds to the membrane structures of tumour cells. Such a ligand can, for example, be attached to the construct or the vector or else be a component of the liposome structure.

15

Suitable ligands are, in particular, polyclonal or monoclonal antibodies, or antibody fragments thereof, which bind, by their variable domains, to the membrane structures of tumour cells, or substances carrying mannose terminally, cytokines or growth factors, or fragments or part sequences thereof, which bind to receptors on tumour cells.

20

Examples of corresponding membrane structures are receptors for a cytokine or a growth factor, such as IL-1, EGF, PDGF, VEGF, TGF β , insulin or insulin-like growth factor (ILGF), or adhesion molecules, such as SLeX, LFA-1, MAC-1, LECAM-1 or VLA-4, or the mannose-6-phosphate receptor.

25

The present invention includes pharmaceutical preparations which, in addition to the vector constructs according to the invention, can also comprise non-toxic, inert, pharmaceutically suitable excipients. It is possible to conceive of administering (e.g. intravenously, intraarterially, intramuscularly, subcutaneously, intradermally, anally, vaginally, nasally, transdermally, intraperitoneally, as an aerosol or orally) these preparations at the site of a tumour or administering them systemically.

30

The vector constructs according to the invention can be employed in gene therapy.

5 The invention furthermore relates to a recombinant host cell, in particular a recombinant eukaryotic host cell, which harbours the above-described constructs or vectors.

10 The invention furthermore relates to a process for identifying substances which affect the promoter activity, silencer activity or enhancer activity of the catalytic telomerase subunit, with this process comprising the following steps:

15 A. adding a candidate substance to a host cell which harbours the regulatory DNA sequence according to the invention, in particular the 5'-flanking regulatory DNA sequence for the gene for the human catalytic telomerase subunit, or a part region thereof which has a regulatory effect, which sequence or part region is functionally linked to a reporter gene, and

B. measuring the effect of the substance on expression of the reporter gene.

20 The process can be employed for identifying substances which increase the promoter activity, silencer activity or enhancer activity of the catalytic telomerase subunit.

25 The process can furthermore be employed for identifying substances which inhibit the promoter activity, silencer activity or enhancer activator of the catalytic telomerase subunit.

30 The invention furthermore relates to a process for identifying factors which bind specifically to fragments of the DNA fragments according to the invention, in particular the 5'-flanking regulatory DNA sequence of the catalytic telomerase subunit. This method comprises screening an expression cDNA library using the above-described DNA sequence, or subfragments of widely differing length, as the probe.

The above-described constructs or vectors can also be used for preparing transgenic animals.

5 The invention furthermore relates to a process for detecting telomerase-associated conditions in a patient, which process comprises the following steps:

10 A. incubating a construct or vector, which contains the DNA sequence according to the invention, in particular the 5'-flanking regulatory DNA sequence for the gene for the human catalytic telomerase subunit, or a part region thereof having a regulatory effect, and a reporter gene, with body fluids or cell samples,

15 B. detecting the activity of the reporter gene in order to obtain a diagnostic value; and

20 C. comparing the diagnostic value with standard values for the reporter gene construct in standardized normal cells or body fluids of the same type as the test sample;

The detection of diagnostic values which are higher or lower than the standard comparative values indicates a telomerase-associated condition, which in turn indicates a pathogenic condition.

25 Explanation of the figures:

Fig. 1: Southern blot analysis using genomic DNA from various species

30 A: Photograph of an ethidium bromide-stained 0.7% agarose gel containing approximately 4 µg of Eco RI-cut genomic DNA. Track 1 contains Hind III-cut λ DNA as size markers (23.5, 9.4, 6.7, 4.4, 2.3, 2.0 and 0.6 kb). Tracks 2 to 10 contain human, rhesus monkey, Sprague

Dawley rat, BALB/c mouse, dog, bovine, rabbit, chicken and yeast (*Saccharomyces cerevisiae*) genomic DNA.

B: Autoradiogram, corresponding to Fig.1 A, of a Southern blot analysis in which radioactively labelled hTC-cDNA probe of about 720 bp in length is used for the hybridization.

Fig. 2: Restriction analysis of the recombinant λ DNA of the phage clone P12, which hybridizes with a probe from the 5' region of the hTC cDNA.

The figure shows a photograph of an ethidium bromide-stained 0.4% agarose gel. Tracks 1 and 2 contain Eco RI/Hind III-cut λ DNA and a 1 kb ladder from Gibco as size markers. Tracks 3 - 7 each contain 250 ng of the DNA from the recombinant phage which has been cut with Bam HI (track 3), Eco RI (track 4), Sal I (track 5), Xho I (track 6) and Sac I (track 7). The arrows mark the two λ arms of the vector EMBL3 Sp6/T7.

Fig. 3: Restriction analysis and Southern blot analysis of the recombinant λ DNA of the phage clone which hybridizes with a probe from the 5' region of the hTC cDNA.

A: The figure shows a photograph of an ethidium bromide-stained 0.8% agarose gel. Tracks 1 and 15 contain a 1 kb ladder from Gibco as size markers. Tracks 2 to 14 each contain 250 ng of cut λ DNA from the recombinant phage clone. The following enzymes were employed: track 2: Sac I, track 3: Xho I, track 4: Xho I, Xba I, track 5: Sac I, Xho I, track 6: Sal I, Xho I, Xba I, track 7: Sac I, Xho I, Xba I, track 8: Sac I, Sal I, Xba I, track 9: Sac I, Sal I, BamH I, track 10: Sac I, Sal I, Xho I, track 11: Not I, track 12: Sma I, track 13: empty, track 14: not digested.

B: Autoradiogram, corresponding to Fig. 3 A, of a Southern blot analysis. A 5'-hTC cDNA fragment of about 420 bp in length was used as the probe for the hybridization.

5 Fig. 4: Partial DNA sequence of the 5'-flanking region and of the promoter of the gene for the human catalytic telomerase subunit. The ATG start codon in the sequence is printed in bold. The depicted sequence corresponds to SEQ ID NO 1.

10 Fig. 5: Use of primer extension analysis to identify the transcription start.

The figure shows an autoradiogram of a denaturing polyacrylamide gel which was selected for depicting a primer extension analysis. An oligonucleotide having the sequence

15 5'GTTAAGTTGTAGCTTACACTGGTTCTC 3' was used as the primer. The primer extension reaction was loaded in track 1. Tracks G, A, T and C constitute the sequence reactions using the same primer and the corresponding dideoxynucleotides. The thick arrow marks the main transcription start while the thin arrows point to three subsidiary transcription start points.

20

Fig. 6: cDNA sequence of the human catalytic telomerase subunit (hTC; cf. our pending application PCT/EP/98/03468). The depicted sequence corresponds to SEQ ID NO 2.

25

Fig. 7: Structural organization and restriction map of the human hTC gene and its 5'-flanking and 3'-flanking regions.

30 Exons are shown as consecutively numbered rectangles which are filled-in in black, and introns are shown as regions which are not filled in. Untranslated sequence segments in the exons are hatched. Translation starts in exon 1 and ends in exon 16. Restriction enzyme cleavage sites

are marked as follows: S, SacI; X, XhoI. The relative arrangement of the five phage clones (P2, P3, P5, P12, P17), and of the product from the genome walking, are shown by thin lines. As the dots indicate, the sequence of intron 16 has only been partly deciphered.

5

Fig. 8: HTL splice variants.

A: Diagrammatic structure of the hTC mRNA splice variants. The complete hTC mRNA is depicted as a rectangle with a grey background in the upper region of the figure. The 16 exons are depicted in accordance with their size. The translation start (ATG) and the stop codon, and also the telomerase-specific T motif, and the seven RT motifs, are all shown. The hTC variants are subdivided into deletion and insertion variants. The missing exon sequences are marked in the deletions. The insertions are shown by additional white rectangles. The sizes and origins of the inserted sequences are given. Newly formed stop codons are marked. The size of the insertion in variant INS2 is unknown.

10

15

B: Exon-intron transitions in the hTC splice variants. Unspliced 5'-flanking and 3'-flanking sequences are shown as white rectangles. The origins of the exon and intron sequences are given. Intron and exon sequences are shown in small letters and large letters, respectively. The donor and acceptor sequences in the splice sites are underlaid as grey rectangles, and their exon and intron origins are also given.

20

25

Fig. 9: Identification of the transcription start by means of RT-PCR analysis.

The RT-PCR was carried out using a cDNA library prepared from HL 60 cells and genomic DNA as the positive control. A common 3' primer hybridizes to a region of the exon 1 sequence. The positions of the different 5' primers in the coding region or the 5'-flanking region are given. In the negative control, no template DNA was added to the PCR reaction. M: DNA size marker.

30

Fig. 10: Nucleotide sequence and structural features of the hTC promoter.

The figure depicts 11273 bp of the 5'-flanking hTC gene sequence, beginning with the translation start codon ATG (+1). The putative region of the translation start is underlined. Possible regulatory sequence segments within the 4000 bp upstream of the translation start are ringed. The depicted sequence corresponds to SEQ ID NO 3.

Fig. 11: Activity of the hTC promoter in HEK-293 cells.

The first 5000 bp of the 5'-flanking hTC gene region are shown diagrammatically in the upper part of the figure. The ATG start codon is picked out. CpG-rich islands are marked by grey rectangles. The sizes of the hTC promoter-luciferase construct are shown on the left-hand side of the figure. The promoterless pGL2 basic construct and the SV40 promoter construct pGL2-Pro were used as controls in each transfection. The relative luciferase activities of the different promoter constructs in HEK cells are shown as continuous bars on the right-hand side of the figure. The standard deviation is indicated. The numerical values represent the average of two independent experiments which were carried out in duplicate.

Tab. 1: Exon-intron transitions in the hTC gene

The table lists the nucleotide sequences at the 3' and 5' splice transitions of the hTC gene. The consensus sequences for donor and acceptor sequences (AG and GT) are underlaid with grey rectangles. The table shows the intron sequences (small letters) and exon sequences (large letters) which flank the splice acceptor and donor sites. The sizes of the exons and introns are given in bp.

Tab. 2: Potential binding sites for DNA-binding factors in the nucleotide sequence of intron 2

The search for possible DNA-binding factors (e.g. transcription factors) was carried out using the "find pattern" algorithm from the Genetics Computer Group (Madison, USA) GCG sequence analysis program package. The table lists the abbreviations of the DNA-binding factors which were identified and their location in intron 2.

Tab. 1

| 3' Acceptor Sequence | | | 5' Donor Sequence | | | | |
|----------------------|----------------------|---------|-----------------------|------|----------------------|----|-------|
| Intron | Exon | Exon bp | Intron | Exon | Intron bp | | |
| | | No. | | | No. | | |
| 5' flanking region | GTTCAGGCAGCGCTGCGT | 1 | CGCCCCCTCCTTCCGCCAG | 281 | gtgggcctccccggggtcg | 1 | 104 |
| cagggcgttccccgcag | GTGTCCTGCCCTGAAGAGC | 2 | TGGCTGCCCAGGAGCCCCAG | 1354 | gtgaggaggtgtgtggccgt | 2 | 8616 |
| catgtccttctcgtttaag | GGGTGGCTGTGTTCCGGC | 3 | TGCAAAAGCATTTGGAATCAG | 196 | gtactgtatccccacgcaca | 3 | 2089 |
| gaggggctctctattgcag | ACAGCACTTGAAGAGGGTG | 4 | GTTCGCCAGAGAAAAGAGG | 181 | gtggctgtgctttgtttta | 4 | 687 |
| cccatgtgtccccgcag | GCCGAGCGTCTCACCTCGA | 5 | TGAGCTGTACTTTGTCAAG | 180 | gtgggtgccggggaccccc | 5 | 494 |
| ctcgctcactcacacag | GTGGATGTGACGGGGCGGT | 6 | CAAGGCTTCAAGAGCCAC | 156 | gttaaggttcacgtgtgata | 6 | >4660 |
| ccctctctctgcgggcag | GTCTCTACCTTGACAGACC | 7 | TGCCGTGTCATCGAGCAG | 96 | gtctgggcactgccctgca | 7 | 980 |
| ctccccgtctgttctgcag | AGCTCCTCCCTGAATGAGG | 8 | CCGTGCCCATCAGGGGCAA | 86 | gtgagtcaggtggccaggt | 8 | 2485 |
| ctgtgtcttccccgcag | GTCTACGTCCAGTGCCAG | 9 | CGGGGATTGGCGGGACGG | 114 | gtgaggcctcctcttcccc | 9 | 1984 |
| gtattttcccttattttag | GCTGCTCCTGCGTTTGGTG | 10 | ACGCGAAAACCTTCCTCAG | 72 | gtgagggccgtgcccgtgtg | 10 | 1871 |
| cattgccccctctgccttag | GACCTGGTCCGAGGTGTC | 11 | TGCAGAGCGACTACTCCAG | 189 | gtgagcgcacctggccgga | 11 | 3801 |
| attccccctgtgtcttag | CTATGCCCGGACCTCCATC | 12 | CCTGTTTCTGGATTTCAG | 127 | gtgagcaggctgatgtca | 12 | 880 |
| tcctttctggcgactcttag | GTGAACAGCCCTCCAGACGG | 13 | TCCTGCTGCAGGCGTACAG | 62 | gtgagccgccaccaagggg | 13 | 3187 |
| ctgtccgcacatcctcttag | GTTTCAGGCATGTGTGCTG | 14 | CTGAAAGCCAAAGAACGCAG | 125 | gtatgtgcaggtgcctggc | 14 | 781 |
| agcctctgttttccccdag | GGATGTCGCTGGGGGCCAA | 15 | CTGGGTCACTCAGGACAG | 138 | gcaagtggtgggtgaggcc | 15 | 536 |
| tctgattttggccccdag | CCCAGACGCAGCTGAGTCG | 16 | TTTTTCAGTTTGAATAAAA | 664 | 3' flanking region | | |

Tab. 2

| Factors | Location in intron 2 |
|-------------------------|--|
| C/EBP | 2925 |
| CRE.2 | 2749 |
| Sp1 | 2378, 4094, 4526, 4787, 4835, 4995 |
| AP-2 CS3 | 5099 |
| AP-2 CS4 | 2213, 3699, 4667, 5878, 5938, 6059, 6180, 6496 |
| AP-2 CS5 | 5350, 5798, 5880, 5940, 6061, 6182, 6375, 6498 |
| PEA3 | 934, 2505 |
| P53 | 2125 |
| GR uteroglobin | 848, 1487, 2956 |
| PR uteroglobin | 3331 |
| Zeste-white | 1577, 1619, 1703, 1745, 1787, 1829, 1871, 1913, 1955, 1997, 2039, 2081, 3518, 3709, 4765, 5014, 5055 |
| GRE | 846 |
| MyoD-MCK right site/rev | 447, 509, 558, 1370, 1595, 1900, 2028, 2099, 4557 |
| MyoD-MCK left site | 108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030 |
| Ets-1 CS | 6408 |
| API | 3784, 4406 |
| CREB | 2801 |
| GATA-1 | 839, 1390, 3154 |
| c-Myc | 108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030 |
| CACCC site | 991 |
| CCAAT site | 1224 |
| CCAC box | 992 |
| CAAT site | 463, 2395 |
| Rb site | 992, 4663 |
| TATA | 3650 |
| CDEI | 106, 1564, 1606, 1690, 1732, 1816, 1900, 1984 |

Examples

The human gene for the catalytic telomerase subunit (ghTC), and the regions of this gene located 5' and 3', were cloned, while the start point for transcription was determined, potential binding sites for DNA-binding proteins were identified and active promoter fragments were highlighted. The sequence of the hTC cDNA (Fig. 6) has already been reported in our application PCT/EP/98/03468, which is also pending. Unless otherwise mentioned, all the data refer to the position of the cDNA in this sequence.

Example 1

A genomic Southern blot analysis was used to determine whether ghTC constitutes a single gene in the human genome or whether there exist several loci for the hTC gene and possibly also ghTC pseudogenes.

In order to do this, a commercially available zoo blot from Clontech was subjected to Southern blot analysis. This blot contains 4 µg of Eco RI-cut genomic DNA from nine different species (human, monkey, rat, mouse, dog, bovine, rabbit, chicken and yeast). With the exception of yeast, chicken and human, the DNA was isolated from kidney tissue. The human genomic DNA was isolated from placenta and the chicken genomic DNA was purified from liver tissue. An hTC cDNA fragment of about 720 bp in length, which was isolated from hTC cDNA, variant Del2 (position 1685 to 2349 plus 2531 to 2590 in Fig. 6 [deletion 2; cf. Example 5 in Fig. 8]), was used as the radioactively labelled probe in the autoradiogram in Fig. 1. The experimental conditions for the blot hybridization and washing steps were taken from Ausubel *et al.* (1987).

In the case of the human DNA, the probe recognizes two specific DNA fragments. The smaller Eco RI fragment, of from about 1.5 to 1.8 kb in length, probably originates from two Eco RI cleavage sites in an intron in the ghTC DNA. On the

basis of this result, it is to be assumed that only one single ghTC gene is present in the human genome.

Example 2

5

In order to isolate the 5' flanking hTC gene sequence, approx. 1.5×10^6 phages from a human genomic placenta gene library (EMBL 3 SP6/T7 from Clontech, order number HL1067j) were hybridized on nitrocellulose filters (0.45 μm ; from Schleicher and Schuell), in accordance with the manufacturer's instructions, with a radioactively labelled 5'-hTC cDNA fragment of about 500 bp in length (position 839 to 1345 in Fig. 6). The nitrocellulose filters were firstly incubated, at 42°C for two hours, in 2 x SSC (0.3 M NaCl; 0.5 M Tris-HCl, pH 8.0) and then in a prehybridization solution (50% formamide; 5 x SSPE, pH 7.4; 5 x Denhard's solution; 0.25% SDS; 100 μg of herring sperm DNA/ml). For the overnight hybridization, the prehybridization solution was supplemented with 1.5×10^6 cpm of denatured, radioactively labelled probe/ml of solution. Nonspecifically bound radioactive DNA was removed under stringent conditions, i.e. by means of three five-minute steps of washing with 2 x SSC; 0.1% SDS at from 55 to 65°C. The filters were evaluated by autoradiography.

20

The phage clones which were identified in this primary investigation were purified (Ausubel *et al.* (1987)). In subsequent analyses, one phage clone, i.e. P12 turned out to be potentially positive. A λ DNA preparation carried out on this phage (Ausubel *et al.* (1987)), and the subsequent restriction digestion with enzymes which release the genomic insert in fragments, showed that this phage clone contains an insert of approx. 15 kb in the vector (Fig. 2).

25

In order to isolate the complete hTC gene sequence, in each case from 1 to 1.5×10^6 phages were screened, in independent experiments, with in each case different radioactively labelled probes, as described above.

30

The phage clones which were identified in these primary investigations, and which were positive for the corresponding probes, were purified. The phage clone P17 was found to contain an hTC cDNA fragment of about 250 bp in length (position 1787 to 2040 in Fig. 6). The phage clone P2 was identified as containing an hTC cDNA
5 fragment of about 740 bp in length (position 1685 to 2349 plus 2531 to 2607 in Fig. 6 [deletion 2; cf. Example 5]). The phage clones P3 and P5 were found to contain a 3' hTC cDNA fragment of 420 bp in length (position 3047 to 3470 in Fig. 6). After the λ DNA had been prepared from these phages, and subsequently subjected to restriction digestion with enzymes which release the genomic insert in fragments, the
10 inserts were subcloned into plasmids (Example 4).

Example 3

In order to investigate whether the 5' end of the hTC cDNA was also present in the
15 insert in the recombinant phage clone P12, the λ DNA from this clone was hybridized, in a Southern blot analysis, with a radiactively labelled hTC cDNA fragment of about 440 bp in length (position 1 to 440 in Fig. 6) from the extreme 5' region (Fig. 3).

20 Since the isolated λ DNA from the positive clone also hybridizes with the extreme 5' end of the hTC cDNA, this phage probably also contains the 5' sequence region flanking the ATG start codon.

Example 4

25 In order to subclone the entire 15 kb insert in the positive phage clone P12 in the form of subfragments, and subsequently to sequence these fragments, restriction endonucleases which, on the one hand, release the entire insert from EMBL3 Sp6/T7 (cf. Example 2) and, in addition, cut within the insert, were selected for digesting the
30 DNA.

In all, two Xho I subfragments, of about 8.3 and about 6.5 kb in length, respectively, and three Sac I subfragments, of about 8.5, about 3.5 and about 3 kb in length, respectively, were subcloned into the pBluescript KS(+) vector (from Stratagene). The 5123 bp 5'-flanking nucleotide sequence of the ghTC gene region, starting from the ATG start codon, was determined by analysing the sequences of these fragments (Fig. 4; corresponding to SEQ ID NO 1). Fig. 4 depicts the first 5123 bp (starting from the ATG start codon). Fig. 10 depicts the entire cloned 5' sequence (corresponding to SEQ ID NO 3).

In order to subclone the entire insert, of approx. 14.6 kb in size, in phage clone P17 in the form of subfragments, restriction endonucleases which, on the one hand, release the entire insert from EMLB3 Sp6/T7 and, in addition, cut a few times within the insert, were selected for digesting the DNA. Three XhoI/BamHI fragments, of 7.1 kb, 4.2 kb and 1.5 kb in size, respectively, and one BamHI fragment, of 1.8 kb in size, were subcloned by means of using a combination digestion with the enzymes XhoI and BamHI. Combination restriction digestion with the enzymes XhoI and XbaI resulted in a XhoI/XbaI fragment of 6.5 kb in size, and two XhoI fragments, of 6.5 kb and 1.5 kb in size, respectively, being cloned.

Digestion with the restriction enzyme XhoI was used to subclone the insert, of approx. 17.9 kb in size, in phage clone P2 in the form of subfragments. In all, three XhoI subfragments, of 7.5 kb, 6.4 kb and 1.6 kb in length, respectively, were cloned. Four SacI fragments, of 4.8 kb, 3 kb, 2 kb and 1.8 kb in size, respectively, were additionally subcloned by digesting with the restriction enzyme SacI.

The insert, of approx. 13.5 kb in size, in phage clone P3 was subcloned by digesting with the restriction enzymes SacI and/or XhoI. Six SacI subfragments, of 3.2 kb, 2 kb, 0.9 kb, 0.8 kb, 0.65 kb and 0.5 kb in length, respectively, and two XhoI subfragments, of 6.5 kb and 4.3 kb in length, respectively, were obtained in this connection.

The insert, of approx. 13.2 kb in size, in phage clone P5 was subcloned by digesting with the restriction enzymes *SacI* and/or *XhoI*. In all, *SacI* fragments of 6.5 kb, 3.3 kb, 3.2 kb, 0.8 kb and 0.3 kb in size, and *XhoI* fragments of 7 kb and 3.2 kb in size, were subcloned.

5

In order to clone the hTC genomic sequence region located 3' of phage clone P17 and 5' of phage clone P2, 3 genomic walkings were carried out using the Clontech GenomeWalker™ kits (catalogue number K1803-1) and various combinations of primers. In a final volume of 50 µl, 10 pmol of dNTP mix were added to 1 µl of human GenomeWalker Library HDL (from Clontech), and a PCR reaction was carried out in 1xKlen Taq PCR reaction buffer and 1xAdvantage Klen Taq polymerase mix (from Clontech). 10 pmol of an internal gene-specific primer, and 10 pmol of the adaptor primer AP1 (5'-GTAATACGACTCACTATAGGGC-3'; from Clontech) were added as primers. The PCR was carried out in 3 steps as a touchdown PCR. First of all, denaturation was carried out at 94°C for 20 sec, and the primers were then annealed, and the DNA chain extended, at 72°C for 4 min, over 7 cycles. There then followed 37 cycles in which the DNA was denaturated at 94°C for 20 sec but the subsequent primer extension took place at 67°C for 4 min. In conclusion, there followed a chain extension at 67°C for 4 min. After this first PCR, the PCR product was diluted 1:50. One µl of this dilution was used in a second nested PCR together with 10 pmol of dNTP mix in 1xKlen Taq PCR reaction buffer and 1xAdvantage Klen Taq polymerase mix and also 10 pmol of a nested gene-specific primer and 10 pmol of the nested Marathon Adaptor primers AP2 (5'-ACTATAGGGCACGCGTGGT-3'; from Clontech). The PCR conditions corresponded to the parameters which were selected in the first PCR. As the sole exception, only 5 cycles rather than 7 cycles were selected in the first PCR step and only 24 cycles, instead of 37 cycles, were run in the second PCR step. The products of this nested genomic walking PCR were cloned into the TA Cloning Vector pCRII from InVitrogen.

30

In the first genomic walking, the gene-specific primer C3K2-GSP1 (5'-GACGTGGCTCTTGAAGGCCTTG-3') and the nested gene-specific primer C3K2-GSP2 (5'-GCCTTCTGGACCACGGCATAACC-3') were used, together with the HDL library 4, and a PCR fragment of 1639 bp in length was obtained. In the second genomic walking, a PCR fragment of 685 bp in length was amplified from the HDL library 4 using the gene-specific primer C3F2 (5'-CGTAGTTGAGCACGCTGAACAGTG-3') and the nested gene-specific primer C3F (5'-CCTTCACCCTCGAGGTGAGACGCT-3. The third genomic walking mixture, using the gene-specific primer DEL5-GSP1 (5'-GGTGGATGTGACGGGCGCGTACG-3') and the nested gene-specific primer C5K-GSP1 (5'-GGTATGCCGTGGTCCAGAAGGC-3'), led to a 924 bp PCR fragments being cloned from the HDL library 1. In all, 2100 bp of the genomic hTC region located 3' of phage clone P17 were identified using this genomic walking method (see Fig. 7).

The subcloned fragments, and the genomic walking products, were sequenced in single-stranded form. The Lasergene Biocomputing Software (DNASTAR Inc. Madison, Wisconsin, USA) was used to identify overlapping regions and form contigs. In all, 2 large contigs were assembled from the sequences collected from phage clones P12, P17, P2, P3 and P5, and also the sequence data from the genomic walking. Contig 1 consists of sequence data from phage clones P12 and P17 and the sequence data from the genomic walking. Contig 2 was put together from the sequences from phage clones P2, P3 and P5. Overlapping phage clone regions are shown diagrammatically in Fig. 7. The sequence data from the 2 contigs are shown below. The ATG start codon in contig 1 is underlined. The TGA stop codon is underlined in contig 2.

Contig1:

| | | | | | | | | |
|----|-------------|-------------|------------|-------------|-------------|-------------|-------------|------|
| | ACTTGAGCCC | AAGAGTTCAA | GGCTACGGTG | AGCCATGATT | GCAACACCAC | ACGCCAGCCT | TGGTGACAGA | 70 |
| 5 | ATGAGACCCT | GTCTCAAAAA | AAAAAAAAAA | AATTGAAATA | ATATAAAGCA | TCTTCTCTGG | CCACAGTGGA | 140 |
| | ACAAACCAG | AAATCAACAA | CAAGAGGAAT | TTTGAAAACT | ATACAAACAC | ATGAAAATTA | AACAATATAC | 210 |
| | TTCTGAATGA | CCAGTGAGTC | AATGAAGAAA | TTAAAAAGGA | AATTGAAAAA | TTTATTTAAG | CAAAATGATA | 280 |
| | CGGAAACATA | ACCTCTCAAA | ACCCACGGTA | TACAGCAAAA | GCAGTGCTAA | GAAGGAAGTT | TATAGCTATA | 350 |
| | AGCAGATACA | TCAAAAAAGT | AGAAAAAGCA | GGCCGAGTGG | CTCATGCCTG | TAATCCCAGC | ACTTTGGGAG | 420 |
| 10 | GCCAAGGCGG | GCAGATCGCC | TGAGGTGAGG | AGTTGCGAGC | CAGCCTGACC | AACACAGAGA | AACCTTGTCG | 490 |
| | CTACTAAAAA | TACAAAATTA | GCTGGGCATG | GTGGCAGATG | CCTGTAATCC | CAGTACTCTG | GGAGGCTGAG | 560 |
| | GCAGGATAAC | CGCTTGAACC | CAGGAGGTGG | AGGTTGCGGT | GAGCCGGGAT | TGCCCCATTG | GACTCCAGCC | 630 |
| | TGGGTAAACA | GAGTGAAACC | CTGTCTCAAG | AAAAAAAAAA | AAGTAGAAAA | ACTTAAAAAT | ACAACCTAAT | 700 |
| | GATGCACCTT | AAAGAAGTAG | AAAAGCAAGA | GCAAACATAA | CCTAAAATTG | GTAAAAGAAA | AGAAAATAA | 770 |
| 15 | AAAGATCAGG | CAGAAATATA | TGAAACTGAA | AGATAACAAT | ACAAAAGATC | AACAAAATTA | AAAGTTGGTT | 840 |
| | TTTTGAAAAG | ATAAACAAAA | TTGACAAACC | TTTGCCACAG | CTAAGAAAAA | AGGAAAGAAG | ACCTAAATAA | 910 |
| | ATAAAGTCAG | AGATGAAAAA | AGAGACATTA | CAACTGATAC | CACAGAAATT | CAAAGGATCA | CTAGAGGCTA | 980 |
| | CTATGAGCAA | CTGTACACTA | ATAAATTGAA | AAACCTAGAA | AAAATAGATA | AATTCCTAGA | TGCATACAAC | 1050 |
| | CTACCAAGAT | TGAACCATGA | AGAAATCCAA | AGCCCAACAA | GACCAATAAC | AATAATGGGA | TTAAAGCCAT | 1120 |
| 20 | AATAAAAAAGT | CTCCTAGCAA | AGAGAAGCCC | AGGACCCAA | GGCTTCCCTG | CTGGATTTTA | CCAATCATTT | 1190 |
| | AAAGAAGAAT | GAATTCCAAT | CCTACTCAAA | CTATTCTGAA | AAATAGAGGA | AAGAATACTT | CCAAACTCAT | 1260 |
| | TCTACATGGC | CAGTATTACC | CTGATTCCAA | AACCAGACAA | AAACACATCA | AAAAACAAAC | AACAAAAAAA | 1330 |
| | CAGAAAGAAA | GAAAACTACA | GGCCAATATC | CCTGATGAAT | ACTGATACAA | AAATCCTCAA | CAAAACACTA | 1400 |
| | GCAAACCCAAA | TTAAACAACA | CCTTCGAAAG | AGATTCCATT | GTGATCAAGT | GGGATTTATT | CCAGGGATGG | 1470 |
| 25 | AAGGATGGTT | CAACATATGC | AAATCAATCA | ATGTGATACA | TCATCCCAAC | AAAAATGAAGT | ACAAAAACTA | 1540 |
| | TATGATTATT | TCACCTTATG | CAGAAAAAGC | ATTTGATAAA | ATTCTGCACC | CTTCATGATA | AAAACCTCTA | 1610 |
| | AAAAACCCGAG | TATACAAGAA | ACATACAGGC | CAGGCACAGT | GGCTCACACC | TGCGATCCCA | GCCTCTGGG | 1680 |
| | AGGCCAAGGT | GGGATGATTG | CTTGGGCCCA | GGAGTTTGAG | ACTAGCCTGG | GCAACAAAAAT | GAGACCTGGT | 1750 |
| | CTACAAAAAA | CTTTTTTAAA | AAATTAGCCA | GGCATGATGG | CATATGCCTG | TAGTCCCAGC | TAGTCTGGAG | 1820 |
| 30 | GCCTGAGGTG | GAGAATCACT | TAAGCCTAGG | AGGTCGAGGC | TGCAGTGAGC | CATGAACATG | TCAGTGTACT | 1890 |
| | CCAGCCTAGA | CAACAGAACA | AGACCCCACT | GAATAAGAAG | AAGGAGAAGG | AGAAGGGAGA | AGGGAGGGAG | 1960 |
| | AAGGGAGGAG | GAGGAGAAGG | AGGAGGTGGA | GGAGAAGTGG | AAGGGGAAGG | GGAAGGGGAA | GAGGAAGAAG | 2030 |
| | AAGAAACATA | TTTCAACATA | ATAAAAGCCC | TATATGACAG | ACCGAGGTAG | TATTATGAGG | AAAAACTGAA | 2100 |
| | AGCCTTTCTCT | CTAAGATCTG | GAAATGACA | AGGGCCCACT | TTCAACCACTG | TGATTCAACA | TAGTACTAGA | 2170 |
| 35 | AGTCCTAGCT | AGAGCAATCA | GATAAGAGAA | AGAAATAAAA | GGCATCCAAA | CTGGAAAGGA | AGAAGTCAAA | 2240 |
| | TTATCCTGTT | TAGCAGATGAT | ATCTGCTTAT | ATCTGGAATA | GACTTAAAGC | ACCACTAAAA | ACTATTGAGT | 2310 |
| | GCTGAAATTT | GGTACAGCAG | GATACAAAAT | CAATGTACAA | AAATCAGTAG | TATTTCTATA | TTCCAACAGC | 2380 |
| | AAACAATCTG | AAAAAGAAAC | CAAAAAAGCA | GCTACAAATA | AAATTAACA | GCTAGGAATT | AACCAAGAA | 2450 |
| | GTGAAGATCT | TCTACAATGA | AAACTATATA | ATGTTGTATA | AAGAAATTGA | AGAGGGACAA | AAAAAGAGAA | 2520 |
| 40 | AGATATTCCA | TGTTTATAGA | TTGGAAGAAT | AAATACTGTT | AAATGTCCA | TACTACCCAA | AGCAATTTAC | 2590 |
| | AAATTCAATG | CAATCCCTAT | TAAATACTA | ATGACGTTCT | TCACAGAAAT | AGAAGAAACA | ATTCTAAGAT | 2660 |
| | TTGTACAGAA | CCACAAAGAG | CCAGAAATAG | CCAAAGCTAT | CCTGACCAAA | AAGAAACAAA | TTGGAAGCAT | 2730 |
| | CACATTACCT | GACTTCAAT | TATACTACAA | AGCTATAGTA | ACCCAAACCTA | CATGGTACTG | GCATAAAAAAC | 2800 |
| | AGATGAGACA | TGGACCAGAG | GAACAGAATA | GAGAATCCAG | AAACAAATCC | ATGCATCTAC | AGTGAACCTA | 2870 |
| 45 | TTTTTGACAA | AGGTGCCAAG | AACATCTTTT | GGGGAAAGAA | TAATCTCTTC | AATAAATGGT | GCTGGAGGAA | 2940 |
| | CTGGATATCC | ATATGCAAAA | TAACAATACT | AGAACTCTGT | CTCTCACCAT | ATACAAAAGC | AAATCAAAAT | 3010 |
| | GGATGAAAGG | CTTAAATCTA | AAACCTCAAA | CTTTGCACT | ACTAAAAGAA | AACACCCGAG | AAACTCTCCA | 3080 |
| | GGACATTTGA | GTGGGCAAG | ACTTCTTGAG | TAATTCCTCTG | CAGGCACAGG | CAACCAAGAG | AAAAACAGAC | 3150 |
| | AAATGGGATC | ATATCAAGTT | AAAAAGCTTC | TGCCAGGCAA | AGGAAACAAT | CAACAAAGAG | AAGAGACAAC | 3220 |
| 50 | CCACAGAATG | GGAGAATATA | TTTGCAAACT | ATTCATCTAA | CAAGGAATTA | ATAACCACTA | TATATAAGGA | 3290 |
| | GCTCAAACTA | CTCTATAAGA | AAAACACCTA | ATAAGCTGAT | TTTCAAAAT | AAGCAAAAGA | CTGGGTGAGA | 3360 |
| | CATTTCTCAA | AAATAAGTCA | ACAAATGGCA | AACAGGCATC | TGAAAATGTG | CTCAACACCA | CTGATCATCA | 3430 |
| | GAGAAATGCA | AATCAAACT | ACTATGAGAG | ATCATCTCAT | CCCAGTTAAA | ATGGCTTTTA | TTCAAAAGAC | 3500 |
| | AGGCAATTAAC | AAATGCCAGT | GAGGATGTGG | ATAAAAGGAA | ACCCTTGGAC | ACTGTTGGTG | GGAAATGAAA | 3570 |
| 55 | TTGCTACCAC | TATGGAGAAC | AGTTTGAAAG | TTCTCAAAA | AACTAAAAAT | AAAGCTACCA | TACAGCAATC | 3640 |
| | CCATTGCTAG | GTATATACTC | CAAAAAAGGG | AATCAGTGTA | TCAACAAGCT | ATCTCCACTC | CCACATTTAC | 3710 |
| | TGCAGCATG | TTCATAGCAG | CCAAGGTTTG | GAAAGCAACT | CAGTGTCAT | CAACAGACGA | ATGGAAAAAG | 3780 |
| | AAAATGTGGT | GCACATACAC | AATGGAGTAC | TACGCAGCCA | TAAAAAAGAA | TGAGATCCTG | TCAGTTGCAA | 3850 |
| | CAGCATGGGG | GGCACTGGTC | AGTATGTTAA | GTGAAATAAG | CCAGGCACAG | AAAGACAAAC | TTTTCATGTT | 3920 |
| 60 | CTCCCTTACT | TGTGGGAGCA | AAAATTAAAA | CAATTGACAT | AGAAATAGAG | GAGAATGGTG | GTCTAGAGG | 3990 |
| | GGTGGGGGAG | AGGGTGACTA | GAGTCAACAA | TAATTTATTG | TATGTTTTAA | AATAACTAAA | AGAGTATAAT | 4060 |
| | TGGGTTGTTT | GTAACACAAA | GAAGGATATA | ATGCTTTGAG | GTGACAGATA | CCCCATTTAC | CCTGATGTGA | 4130 |
| | TTATTACACA | TTGTATGCCT | GTATCAAAAT | ATCTCATGTA | TGCTATAGAT | ATAAACCTTA | CTATATTAAA | 4200 |
| | AATTAATAAT | TTAATGGCCA | GGCACGGTGG | CTCATGTCCG | TAATCCCAGC | ACTTTGGGAG | GCCGAGGGGG | 4270 |
| 65 | GTGGATCACC | TGAGGTGAGG | AGTTTGAAAC | CAGTCTGGCC | ACCATGATGA | AACCCTGTCT | CTACTAAAGA | 4340 |
| | TACAAAATTT | AGCCAGGCGT | GGTGGCACAT | ACCTGTAGTC | CCAACCTACT | AGGAGGCTGA | GACAGAGGAA | 4410 |
| | TGCTTTGAAC | CTGGGAGGCG | GAGGTTGCAG | TGAGCCGAGA | TCATGCCACT | GCACTGCAGC | CTGGGTGACA | 4480 |
| | GAGCAAGACT | CCATCTCAAA | ACAAAAACAA | AAAAAAGAG | ATTAATAATTG | TAATTTTTAT | GTACCGTATA | 4550 |
| 70 | AATATATACT | CTACTATATT | AGAAGTTAAA | AATTAAAAAC | ATTATAAAG | GTAATTAACC | ACTTAATCTA | 4620 |
| | AAATAAGAAC | AATGTATGTG | GGGTTTCTAG | CTTCTGAAGA | AGTAAAAGTT | ATGGCCACGA | TGGCAGAAAT | 4690 |
| | GTGAGGAGGG | AACAGTGGAA | GTTACTGTTG | TTAGACGCTC | ATACTCTCTG | TAAGTGACTT | AATTTTAAAC | 4760 |
| | AAAGACAGGC | TGGGAGAAGT | TAAAGAGGCA | TTCTATAAGC | CCTAAAACAA | CTGCTAATAA | TGGTGAAAGG | 4830 |
| | TAATCTCTAT | TAATTACCAA | TAATTACAGA | TATCTCTAAA | ATCGAGCTGC | AGAATTGGCA | CGTCTGATCA | 4900 |
| | CACCGCTCCT | TCATTACAGG | TGCTTTTTTT | CTTGTGTGCT | TGGAGATTTT | CGATTGTGTG | TTCTGTGTTG | 4970 |
| 75 | GTTAAACTTA | ATCTGTATGA | ATCCTGAAAC | GAAAAATGGT | GGTGATTTC | TCCAGAAGAA | TTAGAGTACC | 5040 |
| | TGGCAGGAAG | CAGGTGGCTC | TGTGGACCTG | AGCCACTTCA | ATCTTCAAGG | GTCTCTGGCC | AAGACCCAGG | 5110 |

| | | | | | | | | |
|----|------------|-------------|------------|-------------|-------------|-------------|-------------|-------|
| | TGCAAGGCAG | AGGCCTGATG | ACCCGAGGAC | AGGAAAGCTC | GGATGGGAAG | GGGCGATGAG | AAGCCTGCCT | 5180 |
| | CGTTGGTGAG | CAGCGCATGA | AGTGCCCTTA | TTTACGCTTT | GCAAAGATTG | CTCTGGATAC | CATCTGGAAA | 5250 |
| | AGGCGGCCAG | CGGGAATGCA | AGGAGTCAGA | AGCCTCCTGC | TCAAACCCAG | GCCAGCAGCT | ATGGCGCCCA | 5320 |
| 5 | CCCGGGCGTG | TGCCAGAGGG | AGAGGAGTCA | AGGCACCTCG | AAGTATGGCT | TAAATCTTTT | TTTACCTGA | 5390 |
| | AGCAGTGACC | AAGGTGTATT | CTGAGGGGAA | CTTGAGTTAG | GTGCTTCTTT | TAAAACAGAA | AGTCATGGAA | 5460 |
| | GCACCTTCT | CAAGGGAATA | CCAGACGCC | GCTCTGCGGT | CATTACCTC | TTTCTCTCT | CCCTCTCTTG | 5530 |
| | CCCTCGCGGT | TTCTGATCGG | GACAGAGTGA | CCCCCGTGA | GCTTCTCCGA | GCCCCGTGCT | AGGACCTCT | 5600 |
| | TGCAAAGGGC | TCCACAGACC | CCCGCCCTGG | AGAGAGGAGT | CTGAGCCTGG | CTTAATAACA | AACCTGGGATG | 5670 |
| 10 | TGGCTGGGGG | CGGACAGCGA | CGGCGGGATT | CAAAGACTTA | ATTCCATGAG | TAAATTC AAC | CTTTCACAT | 5740 |
| | CCGAATGGAT | TTGGATTTTA | TCTTAATATT | TTCTTAAATT | TCATCAAATA | ACATTACAGGA | CTGCAGAAAT | 5810 |
| | CCAAAGGCGT | AAAACAGGAA | CTGAGCTATG | TTTGCCAAAG | TCCAAGGACT | TAATAACCAT | GTTCAGAGGG | 5880 |
| | ATTTTTCGCC | CTAAGTACTT | TTTATTGGTT | TTCTAAGTGT | GGCTTAGGGT | GCAAGGGAAA | GTACACGAGG | 5950 |
| | AGAGGCTTGG | GCGGCAGGGC | TATGAGCAGG | GCAGGGCCAC | CGGGGAGAGA | GTCCCCGGCC | TGGGAGGCTG | 6020 |
| | ACAGCAGGAC | CACGTACCGT | CCTCCCTGGG | AGCTGCCACA | TTGGGCAACG | CGAAGGCGGG | CACGCTGCGT | 6090 |
| 15 | GTGACTCAGG | ACCCCATACC | GGCTTCCTGG | GCCACCCAC | ACTAACCCAG | GAAGTCACGG | AGCTTGAAC | 6160 |
| | CCGTGGAAAC | GAACATGACC | CTTGCTGCTG | TGCTTCCCTG | GGTGGGTCAA | GGGTAATGAA | GTGGTGTGCA | 6230 |
| | GGAAATGGCC | ATGTAATAA | CACGACTCTG | CTGATGGGGA | CCGTTCCTTC | CATCATTTAT | CATCTTCACC | 6300 |
| | CCAAAGGACT | GAATGATTCC | AGCAACTTCT | TCGGGTGTGA | CAAGCCATGA | CAAAACTCAG | TACAAACACC | 6370 |
| | ACTCTTTTAC | TAGGCCACACA | GAGCACGGSC | CACACCCCTG | ATATATTAAG | AGTCCAGGAG | AGATGAGGCT | 6440 |
| 20 | GCTTTTCAGC | ACCAGGCTGG | GGTGACAACA | GCGGCTGAAC | AGTCTGTCTC | TCTAGACTAG | TAGACCTGG | 6510 |
| | CAGGCACTCC | CCCAGATTCT | AGGGCCTGGT | TGCTGTCTCC | CGAGGGCGCC | ATCTGCCCTG | GACTCAGCG | 6580 |
| | CTGGGGTGCC | ACACTGAGGC | CAGCCCTGTC | TCCACACCTT | CCGCTTCCAG | GCCTCAGCTT | CTCCAGCAGC | 6650 |
| | TTCTTAAACC | CTGGGTGGGC | CGTGTTCACG | CGCTACTGTC | TCACCTGTCC | CACGTGTCT | TGTCTCAGCG | 6720 |
| 25 | ACGTAGCTCG | CACGGTTCCT | CCTCACATGG | GGTGTCTGTC | TCCTTCCCCA | ACACTCACAT | CGGTTGAAG | 6790 |
| | GAGGAGATTG | TGCGCCTCCC | AGACTGGCTC | CTCTGAGCCT | GAACCTGGCT | CGTGGCCCCC | GATGCAGGTT | 6860 |
| | CCTGGCGTCC | GGCTGCACGC | TGACCTCCAT | TTCCAGGCGC | TCCCGTCTC | CTGTCACTGT | CCGGGGCCTG | 6930 |
| | CCGGTGTGTT | CTTCTGTGTT | TGTGCTCTTT | TCCAGTCCCA | GCTGCGTGTG | TCTCTGCCCC | CTAGGCTCTC | 7000 |
| | GGGGTTTTTA | TAGGCATAGG | ACGGGGGCGT | GGTGGGCCAG | GGCGCTCTTG | GGAAATGCAA | CATTTGGGTG | 7070 |
| 30 | TGAAAGTAGG | AGTGCCGTGC | CTCACCTTAG | TCCACGGGCA | CAGGCTGGG | GATGGAGCC | CCGCCAGGGA | 7140 |
| | CCCGCCTTCT | TCTGCCAGC | ACTTTCCTGC | CCCCCTCCCT | CTGGAACACA | GAGTGGCAGT | TCCACAAGC | 7210 |
| | ACTAAGCATC | CTCTTCCCAA | AAGACCCAGC | ATTGGCACCC | CTGGACATTT | GCCCCACAGC | CCTGGGAATT | 7280 |
| | CACGTACTGA | CGCACATCAT | GTACACACTC | CCGTCCACGA | CCGACCCCGG | CTGTTTTATT | TTAATAGTGA | 7350 |
| | CAAAGCAGGG | AAATCCCTGC | TAAATGTCTC | TTTAACAAAC | TGGTTAAACA | AACGGGTCCA | TCCGCACGGT | 7420 |
| 35 | GGACAGTTCC | TCACAGTGAA | GAGGAACATG | CCGTTTATAA | AGCCTGCAGG | CATCTCAAGG | GAATTACGCT | 7490 |
| | GAGTCAAAAC | TGCCACCTCC | ATGGGATACG | TACGCAACAT | GCTCAAAAAG | AAAGAATTTT | ACCCCATGGC | 7560 |
| | AGGGGAGTGG | TTAGGGGGGT | TAAGGACGGT | GGGGGGGCGA | GCTGGGGGCT | ACTGCACGCA | CCTTTTACTA | 7630 |
| | AAGCCAGTTT | CTTGTTTCTG | ATGGTATTGG | CTCAGTTATG | GGAGACTAAC | CATAGGGGAG | TGGGGATGGG | 7700 |
| | GGAAACCGGA | GGCTGTGCCA | TCTTTGCCAT | GCCCCAGTGT | CCTGGGCAGG | ATAATGCTCT | AGATGATGCC | 7770 |
| 40 | ACGTCTGTAG | TCCCCAAAC | CTGTGGACAG | AACCCGCGCG | GCCCCAGGGC | CTTTGACAGT | GTGATCTCCG | 7840 |
| | TGAGGACCCT | GAGGTCTGGG | ATCCTTCCGG | ACTACCTGCA | GGCCCCGAAA | GTAATCCAGG | GGTCTTGGA | 7910 |
| | AGAGGCGGGG | AGGAGGGTCA | GAGGGGGGCA | GCCTCAGGAC | GATGGAGGCA | GTCAGTCTGA | GGGTGAAAAG | 7980 |
| | GGAGGGAGGG | CCTCGAGCCC | AGGCCTGCAA | GCGCCTCCAG | AAGCTGGAAA | AAGCGGGGAA | GAGGACCTCC | 8050 |
| | ACGGAGCCTG | CAGCAGGAAG | GCACGGCTGG | CCCTTAGCCC | ACCAGGGCCC | ATCGTGGACC | TCCGGCCTCC | 8120 |
| | GTGCCATAGG | AGGGCACTCG | CGCTGCCCTT | CTAGCATGAA | GTGTGTGGGG | ATTTGAGAGG | GCAACAGGAA | 8190 |
| 45 | ACCCATGCAC | TGTGAATCTA | GGATTATTTC | AAAACAAAGG | TTTACAGAAA | CATCCAAAGG | CAGGGCTGAA | 8260 |
| | GTGCCCTCCG | GCAAGGGCAG | GGCAGGCACG | AGTGATTTTA | TTTAGCTATT | TTATTTTATT | TACTTACTTT | 8330 |
| | CTGAGACAGG | GTTATGCTCT | TGTTGCCAGG | GCTGGAGTGC | AGCGGCATGA | TCTTGCTCTA | CTGCAACCTC | 8400 |
| | CGTCTCCTGG | GTTCAAGCAA | TTCTCGTGCC | TCAGCTCCCC | AAGTAGCTGG | GATTTACAGG | GTGCACACC | 8470 |
| 50 | ACACCCGGCT | AATTTTGTAT | TTTTAGTAGA | GATGGGCTTT | CACCATGTTG | GTCAAGCTGA | TCTCAAAATC | 8540 |
| | CTGACCTCAG | GTGATCCGCC | CACCTCAGCC | TCCCAAAGTG | CTGGGATTAC | AGGCATGAGC | CATGCACTG | 8610 |
| | GGCTATTATT | ACCATTTTAA | AACCTCCCTG | GGCTCAAGTC | ACACCCACTG | GTAAGGAGTT | CTGGAGTTCT | 8680 |
| | AATTTCCCTT | TACTCAGGGA | GTTACCTTCC | TTTGATATTT | TCTGTAATTC | TTCGTAGACT | GGGGATACAC | 8750 |
| | CGTCTCTTGA | CATATTACAC | GTTTCTGTGA | CCACCTGTGA | TCCCATGGGA | CCCAGTGCAG | GGGCAGCTGG | 8820 |
| | GAGGCTGCAG | GGTTCAGGTC | CCAGTGGGGT | TGCCATCTGC | CAGTAGAAAC | CTGATGTAGA | ATCAGGGCGC | 8890 |
| 55 | AAGTGTGGAC | ACTGTCTCTG | ATCTCAATGT | CTCAGTGTGT | GCTGAAACAT | GTAGAAATTA | AAGTCCATCC | 8960 |
| | CTCCTACTCT | ACTGGGATTG | AGCCCCCTCC | CTATCCCCCC | CCAGGGGCAG | AGGAGTTCCT | CTCACTCCTG | 9030 |
| | TGGAGGAAGG | AATGATACTT | TGTTATTTTT | CACTGCTGGT | ACTGAATCCA | CTGTTTCATT | TGTTGGTTTG | 9100 |
| | TTTGTTTTGT | TTTGAGAGGC | GGTTTCACTC | TTGTTGCTCA | GGCTGGAGGG | AGTGCAATGG | CGCGATCTTG | 9170 |
| 60 | GCTTACTGCA | GCCTCTGCCT | CCCAGGTTCA | AGTGATTCTC | CTGCTTCCGC | CTCCCATTTG | GCTGGGATTA | 9240 |
| | CAGGCACCCG | CCACCATGCC | CAGCTAATTT | TTTGATTTTT | TAGTAGAGAC | GGGGGTGGGT | GAGGGTTACC | 9310 |
| | ATGTTGGCCA | GGCTGGTCTC | GAACCTCTGA | CCTCAGATGA | TCCACCTGCC | TCTGCCTCCT | AAAGTGCTGG | 9380 |
| | GATTACAGGT | GTGAGCCACC | ATGCCACGCT | CAGAAATTTAC | TCTGTTTAGA | AACATCTGGG | TCTGAGGTAG | 9450 |
| 65 | GAAGCTCACC | CCACTCAAGT | GTTGTGGTGT | TTTAAGCCAA | TGATAGAATT | TTTTTATTTG | TGTTAGAAAC | 9520 |
| | CTCTTGATGT | TTTACACTGT | GATGACTAAG | ACATCATCAG | CTTTTCAAAG | ACACACTAAC | TGCACCCATA | 9590 |
| | ATACTGGGGT | GTCTTCTGGG | TATCAGCAAT | CTTCATTGAA | TGCCGGGAGG | CGTTTCCCTG | CCATGCACAT | 9660 |
| | GGTGTTAATT | ACTCCAGCAT | AATCTTCTGC | TCTCTTCCCT | TCTCTTCCCT | CTTTTAAAT | CTTTTCTTCT | 9730 |
| | ATGTTGGCTT | CTCTGCAGAG | AACCAAGTGA | AGCTACAAC | TAACCTTTGT | TGGAACAAAT | TTTCCAAACC | 9800 |
| | GCCCCCTTGC | CCTAGTGGCA | GAGACAATTC | ACAAACACAG | CCCTTTAAAA | AGGCTTAGGG | ATCACTAAGG | 9870 |
| 70 | GGATTTCATG | AAGAGCGACC | TGTAATCCTA | AGTATTTACA | AGACGAGGCT | AACCTCCAGC | GACGCTGACA | 9940 |
| | GCCCAGGGAG | GGTGCAGGCG | CTGTTCAAAT | GCTAGCTCCA | TAAATAAAGC | AATTTCTCTC | GGCAGTTTCT | 10010 |
| | GAAAGTAGGA | AAGGTTACAT | TTAAGGTTGC | GTTTGTGTAG | ATTTACAGTGT | TTGCCGACCT | CAGCTCAGCT | 10080 |
| | ATCCCTGCAA | GGCCTCGGGA | GACCCAGAAG | TTTCTCGCCC | CCTTAGATCC | AAACTTGAGC | AACCCGACAG | 10150 |
| | CTGGATTCTT | GGGAAGTCCT | CAGCTGTCTT | GCGGTTGTGC | CGGGGCCCCA | GGTCTGGAGG | GGACCACTGG | 10220 |
| | CCGTGTGGCT | TCTACTGTCT | GGCTGGAAGT | CGGGCTCTCT | AGCTCTGCAG | TCCGAGGCTT | GAGCCAGGCT | 10290 |
| 75 | GCCTGGACCC | CGAGGCTGCC | TCCACCTCTG | TGCGGCGCGG | ATGTGACCAG | ATGTTGGCCT | CTATGCGCAG | 10360 |
| | ACAGAGTGCC | GGGGCCACAG | GTCAAGGCCG | TTGTGGCTGG | TGTGAGGCGC | CCGGTGCCTG | GCCAGCAGGA | 10430 |
| | GCGCTTGGCT | CCATTTCCTA | CCCTTTCTCG | ACGGGACCGC | CCCGGTGGGT | GATTAACAGA | TTTGGGGTGG | 10500 |
| | TTTGCTCATG | GTGGGGACCC | CTCGCCGCCT | GAGAACCTGC | AAAGAGAAAT | GACGGGCTTG | TGTCAGGAG | 10570 |

| | | | | | | | | |
|----|-------------|------------|------------|------------|------------|-------------|------------|-------|
| 5 | CCCAAGTCGC | GGGGAAGTGT | TGCAGGGAGG | CACTCCGGGA | GGTCCCGCGT | GCCCGTCCAG | GGAGCAATGC | 10640 |
| | GTCTCCGGGT | TCGTCCCCAG | CCGCGTCTAC | GCGCCTCCGT | CCTCCCCCTC | ACGTCCGGCA | TTCTGTGGTG | 10710 |
| | CCGGAGCCCG | ACGCCCCCGG | TCCGGACCTG | GAGGCAGCCC | TGGGTCTCCG | GATCAGGCCA | GCGGCCAAG | 10780 |
| | GGTCGCCGCA | CGCACCTGTT | CCCAGGGCCT | CCACATCATG | GCCCCCTCCT | CGGGTTACCC | CACAGCCTAG | 10850 |
| | GCCGATTGCA | CCTCTCTCCG | CTGGGGCCCT | CGCTGGCGTC | CCTGCACCCT | GGGAGCGCGA | GCGGCGCGCG | 10920 |
| | GGCGGGGAAG | CGCGGCCAG | ACCCCCGGGT | CCGCCCGGAG | CAGCTGCCGT | GTGCGGGCCA | GGCCGGGCTC | 10990 |
| | CCAGTGGATT | CGCGGGCACA | GACGCCCAGG | ACCGCGCTCC | CCACGTGGCG | GAGGGACTGG | GGACCCGGGC | 11060 |
| | ACCCGTCTCG | CCCCCTTACC | TTCCAGCTCC | GCCTCCTCCG | CGCGGACCCC | GCCCCGTCCC | GACCCCTCCC | 11130 |
| | GGGTCCCCGG | CCCAGCCCCC | TCCGGGCCCT | CCCAGCCCTT | CCCCCTCCTT | TCCGCGGGCC | CGCCCTCTCC | 11200 |
| 10 | TCGCGGCGCG | AGTTTCAGGC | AGCGCTGCGT | CCTGCTGGCG | ACGTGGGAAG | CCCTGGCCCC | GGCCACCCCC | 11270 |
| | GGGATGCCGC | GGGCTCCCCG | CTGCCGAGCC | GTGCGCTCCC | TGCTGCGCAG | CCACTACCGC | GAGGTGCTGC | 11340 |
| | CGCTGGCCAC | GTTCGTGCGG | CGCCTGGGCG | CCCAGGGCTG | GCGGCTGGTG | CAGCGCGGGG | ACCCGGGCGC | 11410 |
| | TTTCCGCGCG | CTGGTGGCCC | AGTGCCTGGT | GTGCGTGCCC | TGGGACGCAC | GGCCGCCCCC | CGCCGCCCCC | 11480 |
| | TCCTTTCCGCC | AGGTGGGCCT | CCCCGGGGCT | GGCGTCCGGC | TGGGGTTGAG | GGCGGCCGGG | GGGAACCAGC | 11550 |
| 15 | GACATGCCGA | GAGCAGCGCA | GGCGACTCAG | GGCGCTTCCC | CCGAGGTGTG | CCTGCCTGAA | GAGCTGGTGT | 11620 |
| | GCCCGAGTGC | TGCAGAGGCT | GTGCGAGCGC | GGCGCAAGA | ACGTGCTGGC | CTTCGGCTTC | GCGCTGCTGG | 11690 |
| | ACGGGGCCCG | CGGGGGCCCC | CCCCGAGGCT | TCACCACCAG | CGTGCGCAGC | TACCTGCCCA | ACACGGTGAC | 11760 |
| | CGAGCAGCTG | CGGGGAGCG | GGGCGTGGGG | GCTGCTGTGG | GCGCGCTGGG | GCGACAGCGT | GCTGGTTTAC | 11830 |
| 20 | CTGCTGGCAC | GCTGCGCGCT | CTTTGTGCTG | GTGGCTCCCA | GCTGCGCCTA | CCAGGTGTGC | GGGCGCGCGC | 11900 |
| | TGTACCACTG | CGGCGCTGCC | ACTCAGGCCC | GGCCCCCGCC | ACACGCTAGT | GGACCCCGAA | GGCGTCTGGG | 11970 |
| | ATGCGAAGCG | GCCTGGAACC | ATAGCGTCAG | GGAGGCGGGG | GTCCCCCTGG | GCCTGCCAGC | CCCGGGTGGC | 12040 |
| | AGGAGGCGCG | GGGGCAGTGC | CAGCCGAAGT | CTGCCGTTGC | CCAAGAGGCC | CAGGCGTGGC | GCTGCCCTTG | 12110 |
| | AGCCGGAGCG | GACGCCCCGT | GGGCGAGGGT | CCTGGGCCCC | CCCGGGCAGG | ACGCGTGGAG | CGAGTGACCG | 12180 |
| 25 | TGGTTTCTGT | GTGGTGTAC | CTGCCAGACC | CGCCGAAGAA | GCCACCTCTT | TGGAGGTTGC | GCTCTCTGGC | 12250 |
| | ACGCGCCACT | CCCACCCATC | CGTGGGCCGC | CAGCACCACG | CAGGCCCCCC | ATCCACATCG | CGGCCACCAC | 12320 |
| | GTCCCTGGGA | CACGCTTTGT | CCCCCGGTGT | ACGCCGAGAC | CAAGCACTTC | CTCTACTCCT | CAGGCGACAA | 12390 |
| | GGAGCAGCTG | CGGCCCTCCT | TCTACTCAG | CTCTGTGAGG | CCAGCCCTGA | CTGGCGCTCG | GAGGCTCGTG | 12460 |
| | GAGACCATCT | TTCTGGGTTT | CAGGCCCTGG | ATGCCAGGGA | CTCCCCGAG | GTGCCCCGCG | CTGCCCCAGC | 12530 |
| | GCTACTGGCA | AATGCGGCCG | CTGTTTCTGG | AGGTGCTTGG | GAACACGCGG | CAGTGCCCTC | ACGGGGTGGT | 12600 |
| 30 | CCTCAAGACG | CAGTGCCCGC | TGCGAGCTGC | GGTCACCCCA | CGAGCCGGTG | TCTGTGCCCG | CGAGAAGCCC | 12670 |
| | CAGGGCTCTG | TGGCGGCCCC | CGAGGAGGAG | GACACAGACC | CCCGTCGCTT | GGTGCAGCTG | CTCCGCCAGC | 12740 |
| | ACAGCAGCCC | CTGGCAGGTG | TACGGCTTCT | TGCGGGCCGT | CCTGCGCCGG | CTGGTGCCCC | CAGGCCTCTG | 12810 |
| | GGGCTCCAGG | CACAACGAAC | GCCGCTTCCG | CAGGAACACC | AAGAAGTTCA | TCTCCCTGGG | AAGCATGCCC | 12880 |
| 35 | AAGCTCTCGC | TGCAGGAGCT | GACGTGGAAG | ATGAGCGTGC | GGGACTGCGC | TTGGCTGCGC | AGGAGCCAG | 12950 |
| | GTGAGGAGGT | GGTGGCCCTG | GAGGGCCGAG | GCCCCAGAGC | TGAATGCACT | AGGGGCTCAG | AAAAGGGGGC | 13020 |
| | AGGCAGAGCC | CTGGTCTCCT | TGTCTCCATC | GTACGCTGGG | CACACGTGGC | TTTTCTGCTA | CGACGTCGAG | 13090 |
| | TGGACACGGT | GATCTCTGCC | TCTGCTCTCC | CTCTGTCCA | GTTTGCATAA | ACTTACGAGG | TTACACTTCA | 13160 |
| | CGTTTTGATG | GACACGCGGT | TTCCAGGCGC | CGAGGCCAGA | GCAGTGAACA | GAGGAGGCTG | GGCGCGGCAG | 13230 |
| 40 | TGGAGCCGGG | TTGCGGGCAA | TGGGGAGAAG | TGTTTGGAA | CACAGACGCT | CTGGCGAGGG | TGCGTGCAGG | 13300 |
| | TTACCTATAA | TCCTCTTCGC | AATTTCAAGG | GTGGGAATGA | GAGGTGGGGA | CGAGAACCCC | CTCTTCTCTG | 13370 |
| | GGGTGGGAGG | TAAGGGTTTT | GCAGGTGCAC | GTGGTCAGCC | AATATGCAGG | TTTGTGTTTA | AGATTTAATT | 13440 |
| | GTGTGTTGAC | GGCCAGGTGC | GGTGGCTCAC | CCCGGTAATC | CCAGCACTTT | GGGAAGCTGA | GGCAGTTGGA | 13510 |
| | TCACCTGAGG | TCAGGAGTTT | GAGACCAGCC | TGACCAACAT | GGTGAAACCC | TATCTGTACT | AAAAATACAA | 13580 |
| 45 | AAATAGCTG | GGCATGGTGG | TGTGTGCTCG | TAATCCCAGC | TACTTGGGAG | GCTGAGGCGA | GAGAATCACT | 13650 |
| | GAAACTCTGT | CTTTAAAAAA | AAAAAGTGT | CGTTGATTGT | GCCAGGACAG | GGTAGAGGGA | GGGAGATAAG | 13720 |
| | ACTGTTCTCC | AGCACAGATC | CTGGTCCCCA | CTTTAGGTAT | GAAGAGGGCC | ACATGGGAGC | AGAGGACAGC | 13790 |
| | AGATGGCTCC | ACCTGCTGAG | GAAGGGACAG | TGTTTGTGGG | TGTTTGTGGG | ATGGTGCTGC | TGCGCCCTGC | 13860 |
| 50 | CGTGTCCCCA | CCCTGTTTTT | CTGGATTTGA | TGTTGAGGAA | CCTCCGCTCC | AGCCCCCTTT | TGGCTCCGAG | 13930 |
| | TGCTCCCAAG | CCCTACCGTG | GCAGCTAGAA | GAAGTCCCGA | TTTCAACCCC | TCCCCACAAA | CTCCCCAAGC | 14000 |
| | ATGTAAGACT | TCCGGCCATG | CAGACAAGGA | GGGTGACCTT | CTTGGGGCTC | TTTTTTTCTT | TTTCTTCTTT | 14070 |
| | TTATGGTGGC | AAAAGTCATA | TAACATGAGA | TTGGCACTCC | TAACACCGTT | TTCTGTGTAC | AGTGCAGAA | 14140 |
| | TGCTAACTCG | GCGGTGTTTA | CAGCAGGTTG | CTTGAAATGC | TGCGTCTTGC | GTGACTGGAA | GTCCCTACCC | 14210 |
| 55 | ATCGAACCAG | AGCTGCCCTA | CACCTGCTGC | GGCTCAGGTG | GACCAACGCG | AGTCAGATAA | GGCTCATGCA | 14280 |
| | ACCCAGTTTT | GCTTTTTGTG | CTCCAGCTTC | CTTCGTTGAG | GAGAGTTTGA | GTCTCTGAT | CAGGACTCTG | 14350 |
| | CGTGTCAATT | CTGTTCTCTG | ACTTCAGATG | AGGTCACAAT | CTGCCCCCTG | CTTATGCAGG | GAGTGAGGCG | 14420 |
| | TGGTCCCCGG | GTGTCCCTGT | CACGTGCAGG | GTGAGTGAGG | CGTTGCCCCC | AGGTGTCCCT | GTACAGTGTG | 14490 |
| | GGGTGAGTGA | GGCGCGGCC | CCGGGTGTCC | CTGTCCCGTG | CAGCGTGATT | GAGGTGTGGC | CCCCGGGTGT | 14560 |
| 60 | CCCTGTCAAG | TGTAGGGTGA | GTGAGGCGCC | ATCCCCGGGT | GTCCCTGTCA | CGTGTAGGGT | GAGTGAGGCG | 14630 |
| | TGGTCCCCGG | GTGTCCCTGT | CCCGTGCAGG | GTGAGTGAGG | CACGTGCCCC | GGGTGTCCCT | GTACAGTGTG | 14700 |
| | GGGTGAGTGA | GGCGCGGTCC | CCGGGTGTCC | CTCTCAGGTG | TAGGGTGAGT | GAGGCGCGCG | CCCAGGGTGT | 14770 |
| | CCCTGTCAAG | TGTAGGGTGA | GTGAGGACCA | GTCCCTGGGT | GTCCCTCCCC | GGTATAGGGT | GAGTGAGGCA | 14840 |
| | CTGTCCCCGG | GTGTCCCTGT | CACGTGCAGG | GTGAGTGAGG | GCGGGCCCCC | GGGTGTCCCT | CTCAGGTGCA | 14910 |
| 65 | GGGTGAGTGA | GGCGCTGTCC | CTGGGTGTCC | CTGTCTCTGT | TAGGGTGAGT | GAGGCTCTGT | CCCCAGGTGT | 14980 |
| | CCTTGGCGTT | TGCTCACTTG | AGCTTGTCTC | TGAATGTTTG | CTCTTTCTAT | AGCCACAGCT | GCGCCGGTTG | 15050 |
| | CCCATGGCCT | GGGTAGATGG | TGCAGGCCGA | TGCTGTGTTT | CCAAGCCTAT | CTTTCTGAT | GTCGGCTCT | 15120 |
| | TCTTGSTCAC | CTCTCCGTTT | CATTTTGTCT | CGGGACACAG | GGACTGCAGG | CTCTCGCTC | CCGCGTGCCA | 15190 |
| | GGCATGAGAG | CCACAGCTTC | AGGTCCGCTT | GCCTCTGTTG | GGCCTGGCTT | GCTCACCACG | TGCCCCGCCA | 15260 |
| 70 | ATGCAATGCT | CCAATACTCC | TCTCCAGCT | TGTTCCAGCT | CGAGGCTGGA | CTCTGGGCTG | CTGTGTCTG | 15330 |
| | CTGCCACGTG | TTGCTGGAGA | CATCCAGGAA | AGGGTCTCTT | GTGCCCTGAA | GGAAAGCAAG | TCACCCAGC | 15400 |
| | CCCTCTCACT | GTCTGTTTTT | CTCCCCAGCT | GCCCCCTGCG | TTGGCCCCCT | TGGGTGGGTG | GCAACGCTTG | 15470 |
| | TCACCTTATT | CTGGGCACCT | GCGGCTCATT | GCTTAGGCTG | GGCTCTGCTT | CCAGTCGCCC | CCTCATATGG | 15540 |
| | ATTGACGTCC | AGCCACAGGT | TGGAGTGTCT | CTGTCTGTCT | CCTGCTCTGA | GACCCACGTG | GAGGGCCGGT | 15610 |
| 75 | GTCTCCGCGA | GCCTTCGTC | GACTTCCCTC | TTGGGTCTTA | GTTTTGAATT | TCACTGATTT | ACCTCTGACG | 15680 |
| | TTTCTATCTC | TCCATTGTAT | GCTTTTTCTT | GGTTTATCTT | TTTCTTCTTT | TTTCTGATTT | TGAGTTTATG | 15750 |
| | CATGCCCTTC | CCTCTAAGTG | CTGCCCTTAC | TGCACCTGCT | GTTTTGATGT | GAAGTAATCT | CAACATCAGC | 15820 |
| | CACCTTCAAG | TGTTCTTAAA | ATACTTCAAA | GTGTAAATAC | TTCTTTTAA | TATTTCTTAT | CTGTGATTTT | 15890 |
| | TTTCTTTGTG | CACGCTGTGT | TTTGACGTGA | AATCATTTTG | ATATCAGTGA | CTTTTAAAGTA | TTCTTTAGCT | 15960 |

| | | | | | | | | |
|----|------------|-------------|------------|-------------|------------|-------------|-------------|-------|
| | TATTCTGTGA | TTTCTTTGAG | CAGTGAGTTA | TTTGAACACT | GTTTATGTTT | AAGATATGTA | GAGTATCAAG | 16100 |
| | ATACGTAGAG | TATTTTAAAGT | TATCATTTTA | TTATTGATTT | CTAACTCAGT | TGTGTAGTGG | TCTGTATAAT | 16170 |
| | ACCAATTATT | TGAAGTTTGC | GGAGCCTTGC | TTTGTGATCT | AGTGTGTGCA | TGGTTTCCAG | AACTGTCCAT | 16240 |
| 5 | TGTAAATTTG | ACATCCTGTC | AATAGTGGGC | ATGCATGTTT | ACTATATCCA | GCTTATTAAG | GTCCAGTGCA | 16310 |
| | AAGCTTCTGT | CTCCTTCTAG | ATGCATGAAA | TTCCAAGAAG | GAGGCCATAG | TCCCTCACCT | GGGGGATGGG | 16380 |
| | TCTGTTTATT | TCTTCTCGTT | TGGTAGCATT | TATGTGAGGC | ATTGTTAGGT | GCATGCACGT | GGTAGAATTT | 16450 |
| | TTATCTTCCT | GATGAGTGAA | TCTTTTGGAG | ACTTCTATGT | CTCTAGTAAT | CTAGTAATTC | TTTTTTTAAA | 16520 |
| | TTGCTCTTAG | TACTGCCACA | CTGGGCTTCT | TTTGATTAGT | ATTTTCTGCT | TGTGTCTGTT | TTCTGCCTTT | 16590 |
| 10 | AATTTATATA | TATATATATA | TTTTTTTTTT | TTTTGAGACA | GAGTCTTGGT | CTGTGCGCCA | GGGTGAGTGC | 16660 |
| | AGTGGTGTGA | TCACAGGTCA | GTGTAACCTT | TACCTTCTGG | CCTGAGCCGT | CCTCTCACCT | CAGCCTCCTG | 16730 |
| | AGTAGCTGGA | ACTGCAGACA | CGCACCGCTA | CACCTGGCTA | ATTTTAAAT | TTTTTCTGGA | GACAGGGTCT | 16800 |
| | TGGCGTGTG | CCCAGGCTGG | TCTCAAACCT | TTGGACTCAA | GGGATCCATC | TACCTCGGCT | TCCCAAAGTG | 16870 |
| | CTGAATTACA | GGCATGAGCC | ACCATGTCTG | GCCTAATTTT | CAACACTTTT | ATATTCTTAT | AGTGTGGGTA | 16940 |
| | TGTCCTGTGA | ACAGCATGTA | GGTGAATTTT | CAATCCAGTC | TGACAGTCGT | TGTTAACTG | GATAACCTGA | 17010 |
| 15 | TTTATTTTCA | TTTTTTTGTG | ACTAGAGACC | CGCCTGGTGG | ACTCTGATTG | TCCACTTGCC | GTTTGCATCT | 17080 |
| | CCTCGTTCCT | TTGTTTCTCA | CCACCTCTTG | GGTTGCCATG | TGCGTTTCTT | GCCGAGTGTG | TGTTGATCCT | 17150 |
| | CTCCTTGCTT | CCTGGTCACT | GGGCATTTGC | TTTTATTCTT | CTTTGCTTAG | TGTTACCCCT | TGATCTTTTT | 17220 |
| | ATTGTCGGTG | TTTGCTTTTG | TTTATTGAGA | CAGTGTCACT | CTGTCACCCA | GGCTGGAGTG | TAATGTCACA | 17290 |
| 20 | ATCTCGGCTC | ACTGCAACCT | CTGCCCTCCT | GGTTCAAGCA | GTTTCTCATT | CTCAACCTCA | TGAGTAGCTG | 17360 |
| | GGATTACAGG | CGCCACCAC | CACGCGCTGG | TAATTTTGTG | ATTTTTAGTA | GAGATAGGCT | TTACCATATG | 17430 |
| | TGGCCAGGCT | GGTCTCAAAC | TCTGACCTC | AAGTAGTCA | CCCGCTTGG | CCTCCACAG | TCTGGGAT | 17500 |
| | ACAGGTGCAA | GCCACCGTGC | CCGGCATACC | TTGATCTTTT | AAAATGAAGT | CTGAAACATT | GCTACCCCTG | 17570 |
| | TCCTGAGCAA | TAAGACCTTT | AGTGTATTTT | AGCTCTGGCC | ACCCCCACAG | CTGTGTGCTG | TTTTCCCTGC | 17640 |
| 25 | TGACTTAGTT | CTATCTCAGG | CATCTTGACA | CCCCACAAG | CTAAGCATTG | TTAATATTGT | TTTCCGTGTT | 17710 |
| | GAGTGTCTTT | GTAGCTTTGC | CCCCGCCCTG | CTTTTCTCTC | TTTGTTCCTC | GTCTGTCTTC | TGTTCTCAGG | 17780 |
| | CCGCGCTGCT | GGGTCCCTTT | CCTTGTCTCT | TGCGTGGTTC | TTCTGTCTTG | TTATTGCTGG | TAAACCCACG | 17850 |
| | CTTTACCTGT | GCTGGCTCTC | ATGGCATCTA | GCGAGCTCCG | GGGACCTCTG | CTTATGATGC | ACAGATGAAG | 17920 |
| | ATGTGGAGAC | TCACGAGGAG | GGCGGTCTAT | TTGGCCCTGG | AGTGTCTGGA | GCACCCAGTG | GCCAGCGTTC | 17990 |
| 30 | CTTAGCCAGT | GAGTGACAGT | AACGTCCGCT | CGGCTCGGTT | TCAGCCTGGA | AAACCCACAG | CATGTGCGGG | 18060 |
| | TCTGGTGGCT | CCGCGGTGTC | GAGTTTGAAA | TCGCGCAAA | CTGCGGTGTC | GCGCCAGCTC | TGACGGTGCT | 18130 |
| | GCCTGGCGGG | GGAGTGTCTG | CTTCTCTCCT | TCTGCTTGGG | AACCAGGACA | AAGGATGAGG | CTCCGAGGCT | 18200 |
| | TTGTGCGCCA | ACAGGAGCAT | GACCTGAGCC | ATGTGGATAA | TTTTAAATTT | TCTAGGCTGG | GCGCGGTGGC | 18270 |
| | TCACGCTGCT | AATCCCAGCA | CTTTGGGAGG | CCAAGGCGGG | TGGATCACGA | GGTCAGGAGG | TCCGAGCAT | 18340 |
| 35 | CCTGGCCAA | ATGATGAAAC | CCCATCTGTA | CTAAAAACAC | AAAAATTAGC | TGGCGGTGGT | TCGCGGTGCC | 18410 |
| | TGTAATCCCA | GCTACTCGGG | AGGCTGAGGC | AGGAGAAATG | CTTGAACCTG | GGAGTTGGAA | GTTGCAAGTA | 18480 |
| | GCCGACATTG | CACCACTGCA | CTCCAGCTCG | GCAACACAGC | GAGACTCTGT | CTCAAAATAA | AAAAAAATAA | 18550 |
| | AAAAAAATAA | AATTCTAGTA | GCCACATTAA | AAAAATAAAA | AAGAAAGGTT | GAAATTAATG | TAATAATAGA | 18620 |
| | TTTTACTGAA | GCCACAGCAT | TCCACACCTC | ATCATTTTAT | GGTGTATTG | GTGGGAGCAT | CACCTCACAG | 18690 |
| 40 | ACATTGACA | TTTTTTGAGC | TTTGTCTGCG | GGATCCCCTG | TGTAGGTCCC | GTGCGTGGCC | ATCTCGGCTG | 18760 |
| | GGACCTGCTG | GGCTTCCCAT | GGCCTAGGCT | GTTGTACCAG | ATGGTGCAGG | TCCGGGATGA | GGTCGCCAGG | 18830 |
| | CCCTCAGTGA | GCTGGATGTC | CAGTGTCCGG | ATGGTGCAGC | TCTGGGATGA | GGTCGCCAGG | CCCTGCTGTT | 18900 |
| | AGCTGAGTGT | TGCTGTCTG | GATGGTGCAG | GTCAGGGGTC | AGGCTCCAG | GCCCTCGGTG | AGCTGGAGTG | 18970 |
| | ATGGAGTCCG | GATGATGCAG | GTCGGGGGTC | AGGTCGCCAG | GCCCTGCTGT | GAGCTGGATG | TGTGGTGTCT | 19040 |
| 45 | GGATGGTGCA | GGTCAGGGGT | GAGGTCTCCA | GGCCCTCGGT | AAGCTGGAGG | TATGGAGTCC | GGATGATGCA | 19110 |
| | GGTCCGGGGT | GAGGTCCGCA | GGCCCTGCTG | TGAGCTGGAT | GTGTGGTGT | TGGATGGTGC | AGGTCTGGGG | 19180 |
| | TGAGGTACCC | AGGCCCTGCG | GTGAGCTGGG | TGTGCGGTGT | CTGGATGGTG | CAGGTCTGGA | GTGAGGTCCG | 19250 |
| | CAGACGGTGC | CAGACCATGC | GGTGAGCTGG | ATATGCGGGT | TCCGGATGTT | GCAGGTCTGG | GGTGAGGTTG | 19320 |
| | CCAGGCCCTG | CTGTGAGTGT | GATGTGGGGT | GTCCGGATGC | TGCAGGTCCG | GTGTGAGGTC | ACCAGGCCCT | 19390 |
| 50 | GCTGTGAGCT | GGATGTGTGG | TGTCTGGATG | GTGCAGGTCT | GGGGTGAAGG | TGCGCCAGGCC | CCTGCTTGTG | 19460 |
| | AGCTGGATGT | TGTTGTGCTG | GATGGTGCAG | GTCTGGAGAT | AGGTCGCCAG | GCCCTCGGTC | AGCTGGATGT | 19530 |
| | GCAGTGTCCA | GATGGTGCAG | GTCCGGGGTG | AGGTCGCCAG | ACCTGCGGTC | GAGCTGGATG | TGCGGTGTCT | 19600 |
| | GGATGGTGCA | GGTCTGGAGT | GAGGTCCGCA | GGCCCTCGGT | GAGCTGGATG | TATGGAGTCC | GGATGGTGCC | 19670 |
| | GGTCCGGGGT | GAGGTCCGCA | GACCTCTGCT | TGAGCTGGAT | GTGCGGTGTC | TGGATGGTAC | AGGTCGGAG | 19740 |
| 55 | TGAGGTCCGC | AGACCTGTCT | GTGAGCTGGA | TATGCGGTGT | CCGGATGGTG | CAGGTCAAGG | GTGAGGTCTC | 19810 |
| | CAGGCCCTCG | GTGAGCTGGA | GGTATGGAGT | CCGGATGATG | CAGGTCCGGG | GTGAGGTCCG | CAGGCCCTGC | 19880 |
| | TGTGAACCTG | ATGTGCGGGC | TCTGGATGTT | GCAGGTCTGG | GGTGTGGTGC | CCAGGCCCTC | GGTGAGCTGG | 19950 |
| | AGGTATGGAG | TCCGGATGAT | GCAGGTCCGG | GGTGAGGTCG | CCAGGCCCTG | CTGTGAGCTG | GATGTGCGGC | 20020 |
| | GCTCTGGATG | TGCAGGTCTG | GGGTGTGGTC | GCCAGGCCCT | CGGTGAGCTG | GAGGTATGGA | GTCCGGATGA | 20090 |
| 60 | TGCAGGTCCG | GGGTGAGGTT | GCCAGGCCCT | GCTGTGAGCT | GGATGTGCTG | TATCCGGATG | GTGCAGTCCG | 20160 |
| | GGGTGAGGTC | GCCAGGCCCT | GCTGTGAGCT | GATGTGCTG | TATCCGGATG | GTGCAGGTC | GGGTGAGGTC | 20230 |
| | CACCAGGCC | TGCGGTGAGC | TGGTTGTGCG | GTGTCCGGTT | GCTGCAGGTC | CGGGGTGAGT | TGCGCCAGGCC | 20300 |
| | CTCGGTGAGC | TGGATGTGCG | GTGTCCCGCT | GTCCGGATGC | TGCAGGTCCA | GGGTGAGGTC | GCTAGGCCCT | 20370 |
| 65 | TGGTGGGCTG | GATGTGCGGT | GTCCGGATGG | TGCAGGCTG | GGGTGAGGTC | GCCAGGCCCT | TGGTGGGCTG | 20440 |
| | GATGTGCGGT | GTCTGCATGG | TGCAGGTCTG | GGGTGAGGTC | GCCAGGCCCT | TGGTGGGCTG | GATGTGTGGT | 20510 |
| | GTCCGGATGG | TGCAGGTCCG | CGGTGAGGTC | GCCAGGCCCT | GCTGTGAGCT | GGATGTGCGG | TGTCTGGATG | 20580 |
| | GTGCAAGTCC | GGGTGAGGTC | AGCCAGGCC | TTGGATGTGG | GTGTCCGGAT | GTGTCCGGAT | GTTGCAAGTC | 20650 |
| | CGGGGTGAGG | TGCGCCAGGCC | CTGCGGTTAG | CTGGATATGC | GGTGTCCGGA | TGGTGCAGGT | CCGGGGTGAG | 20720 |
| | GTCAACAGGC | CCTGCGGTTA | GCTGGATGTC | CGGTGTCTAG | ATGGTGCAGG | TCCGGGGTGA | GGTCGCCAGG | 20790 |
| 70 | CCCTGCTGTG | AGCTGGATGT | GCTGTATCCG | GATGGTGCAG | GATGGTGCAG | AGGTGCGCAG | GCCCTGCATG | 20860 |
| | GAGCTGGATG | TGCTGTATCC | GGATGGTGCA | GGTCTGGCGT | GAGGTGCGCA | GGCCCTGCGG | TTAGCTGGAT | 20930 |
| | ATCGGGTGTG | GGATGGTGCA | GGTCCGGGCT | GAGGTCAACCA | GGCCCTGCGG | TTAGCTGGAT | GTGCGGTGTC | 21000 |
| | CGGATGGTGT | AGGTCTGGGG | TGAGGTCCGC | AGGCCCTGCT | GTGAGCTGGA | TGTGCTGTAT | CCGGATGGTG | 21070 |
| | CAGGTCCGGG | GTGAGGTCCG | CAGGCCCTGC | GGTGAGCTGG | ATGTGCTGTA | TCCGGATGGT | GCAGGTCTGG | 21140 |
| 75 | CGTGAGGTGC | CCAGGCCCTG | CGGTGAGCTG | GATGTGCACT | GTACGGATGG | TGCAGGTCCG | GGGTGAGGTC | 21210 |
| | GCCAGGCCCT | CGGTGGGGCT | GTATGTGTGT | TGTCTGGATG | GTGCAGGTCC | GGGGTGGAGT | GCCAGGCCCT | 21280 |
| | TGCGGTGAGC | TGGATGTGTG | GTGTCTGGAT | GCTGCAGGTC | CGGGGTGAGT | TGCGCCAGGCC | CTCGGTGAGC | 21350 |
| | TGGATATGCG | GTGTCCCGCT | GTCCGAATGG | TGCAGGTCCA | GGGTGAGGTC | GCCAGGCCCT | TGGTGGGCTG | 21420 |
| | GATGTGCCGT | GTCCGGATGG | TGCAGGTCTG | GGGTGAGGTC | GCCAGGCCCT | TGGTGGGCTG | GATGTGCGGT | 21490 |

| | | | | | | | | |
|----|-------------|-------------|------------|-------------|------------|-------------|-------------|-------|
| | GTCCGGATGG | TGCAGGTCCG | GGGTGAGGTC | ACCAGGCCCT | CGGTGATCTG | GATGTGGCAT | GTCTTCTCG | 21560 |
| | TTTAAGGGGT | TGGCTGTGTT | CCGGCCGAG | AGCACCCTCT | CGGTGAGGAG | ATCCTGGCCA | AGTTCTTGCA | 21630 |
| | CTGGCTGATG | AGTGTGTACG | TCGTCGAGCT | GCTCAGGTCT | TTCTTTTATG | TCACGGAGAC | CACGTTTCAA | 21700 |
| 5 | AAGAACAGGC | TCTTTTCTA | CCGGAAGAGT | GTCTGGAGCA | AGTTGCAAAG | CATTGGAATC | AGGTACTGTA | 21770 |
| | TCCCCACGCC | AGGCCTCTGC | TTCTCGAAGT | CCTGGAACAC | CAGCCCGGCC | TCAGCATGCG | CCTGTCTCCA | 21840 |
| | CTTGCCTGTG | CTTCCCTGGC | TGTGCACTC | TGGGCTGGGA | GCCAGGGGCC | CCGTCACAGG | CCTGGTCCAA | 21910 |
| | GTGGATTCTG | TGCAAGGCTC | TGACTGCCTG | GAGCTCACGT | TCTCTTACTT | GTAATAATCAG | GAGTTTGTGC | 21980 |
| | CAAGTGGTCT | CTAGGGTTTG | TAAAGCAGAA | GGGATTTAAA | TTAGATGGAA | ACACTACCAC | TAGCCTCCTT | 22050 |
| 10 | GCCTTTCCCT | GGGATGTGGG | TCTGATTCTC | TCTCTCTTTT | TTTTTTCTTT | TTTGAGATGG | AGTCTCACTC | 22120 |
| | TGTTGCCAG | GCTGGAGTGC | AGTGGCATAA | TCTTGGCTCA | CTGCAACCTC | CACCTCCTGG | GTTTAAGCGA | 22190 |
| | TTCAACAGCC | TCAGCCTCCT | AAGTAGCTGG | GATTACAGGC | ACCTGCCACC | ACGCCTGGCT | AATTTTTTGT | 22260 |
| | CTTTTAGGAG | AGACGGGGTT | TCACCATGTT | GGCCAGGGTG | GCTCTGAACT | CATGACCTCA | GGTGATCCAC | 22330 |
| | CCACCTTGGC | CTCCCAAAGT | GCTGGGTTTA | CAGGCTAAGC | CACCGTGCCC | AGCCCCGAT | TCTCTTTTAA | 22400 |
| 15 | TTACATGCTGT | TCTGTATGAA | TCTTCAATCT | ATTGGATTTA | GGTCATGAGA | GGATAAAATC | CCACCCACTT | 22470 |
| | GGCGCATGAT | TGCAGGGAGC | ACCTGTGCAG | GGAGCACTG | GGGATAGGAG | AGTTCCACCA | TAGCTAACT | 22540 |
| | TCTAGGTGGC | TGCATTTGAA | TGGCTGTGAG | ATTTTGTCTG | CAATGTTCCG | CTGATGAGAG | TGTGAGATTG | 22610 |
| | TGACAGATTCT | AACTGGGATT | TGCATCAGTG | AGGGACGGGA | GCCTGGTCT | GGGAGATGCC | AGCCTGGCTG | 22680 |
| | AGCCAGGCC | ATGGTATTAG | CTTCTCCCTG | TCCCGCCAG | GCTGACTGTG | GAGGGCTTTA | TACCGTAACT | 22750 |
| 20 | CAGGGCTTCC | CCAGCTCCCC | TGCACACTCG | AGTCCCTGGG | GGGCTTGTG | ACACCCCATG | CCCCAAATCA | 22820 |
| | GGATGTCTGC | AGAGGGAGCT | GGCAGCAGAC | CTCGTCAGAG | GTAACACAGC | CTCTGGGCTG | GGGACCCCGA | 22890 |
| | CGTGGTGTG | GGGCCATTTC | CTTGCATCTG | GGGGAGGGTG | AGGGCTTTTC | CTGTGGGAAC | AGTTTAATAC | 22960 |
| | ACAATGCACC | TTACTTAGAC | TTTACACGTA | TTTAATGGTG | TGCGACCCAA | CATGGTCAAT | TGACCAAGTAT | 23030 |
| | TTTGGAAAAG | ATTTAATTGG | GGTGACCGGA | AGGAGCAGAC | AGACGTGGTG | GTCCCCAAGA | TGCTCCTTGT | 23100 |
| 25 | CACACTAGTG | ACTGTTGTTT | TGCTTGGGGT | GCCTTGGAGG | CCCTCCTCTC | CTGGACAGGG | TACCGTCCCT | 23170 |
| | TTTCTACTCT | GCTGGGCCCTG | CGGCCCTGCG | TCAGGGCACC | AGCTCCGGAG | CACCCGCGGC | CCCAGTGTCC | 23240 |
| | ACGGAGTGCC | AGGCTGTGAG | CCACAGATGC | CCAGGTCAG | GTGTGGCCGC | TCCAGCCCCC | GTGCCCCCAT | 23310 |
| | GGGTGGTTTT | GGGGAAAAG | GCCAAGGGCA | GAGGTGTGAG | GAGACTGGTG | GGCTCATGAG | AGCTGATTCT | 23380 |
| | GCTCCTTGGC | TGAGCTGCCC | TGAGCAGCCT | CTCCCGCCCT | CTCCATCTGA | AGGGATGTGG | CTCTTTCTAC | 23450 |
| | CTGGGGTGCC | TGCTTGGGGC | CAGCCTTGGG | CTACCCCATG | GGCTGTACCA | GAGGGACAGG | CATCCTGTGT | 23520 |
| 30 | GGAGGGCAT | GGGTTACAGT | GGCCCCAGAT | GCAGCGCTGG | ACCAGGCTCC | CTGGTGCTGA | TGGTGGGACA | 23590 |
| | GTCACCCCTG | GGGTTGACCG | CCGGACTGGG | CGTCCCCAGG | GTGACTATA | GGACCAAGTG | TCCAGGTGCC | 23660 |
| | CTGCAAGTAG | AGGGGCTCTC | AGAGGCGTCT | GGCTGGCATG | GGTGGACGTG | GCCCCGGGCA | TGGCCTTCAG | 23730 |
| | CGTGTGCTGC | CGTGGGTGGC | CTGAGCCCTC | ACTGAGTCTG | TGGGGGCTTG | TGGCTTCCCG | TGAGCTTCCC | 23800 |
| | CCTAGTCTGT | TGTCTGGCTG | AGCAAGCCTC | CTGAGGGGCT | CTCTATTGCA | GACAGCACTT | GAAGAGGGTG | 23870 |
| 35 | CAGCTGCGGG | AGCTGTGCGA | AGCAGAGGTC | AGGCAGCATC | GGGAAGCCAG | GCCCCGCTGT | CTGACGTCCA | 23940 |
| | GACTCCGGCT | CATCCCCAAG | CCTGACGGGC | TGCGGCCGAT | TGTGAACATG | GACTACCTCG | TGGGAGCCAG | 24010 |
| | AACGTTCCGC | AGAGAAAAGA | GGGTGGCTGT | GCTTTGGTTT | AACTTCTCTT | TTAAACAGAA | GTGCGTTTGA | 24080 |
| | GCCCCACATT | TGGTATCAGC | TTAGATGAAG | GGCCCGGAGG | AGGGGCCACG | GGACACAGCC | AGGGCCATGG | 24150 |
| 40 | CACGGGCGCA | ACCCATTGTT | GGCAGACATG | AGGTGGCCGA | GGTGCCGGTG | CCTCCAGAAA | AGCAGCGTGG | 24220 |
| | GGGTGTAGGG | GGAGCTCCTG | GGGCAGGGAC | AGGCTCTGAG | GACCACAAGA | AGCAGCCGGG | CCAGGGCCTG | 24290 |
| | GATGCAGCAC | GGCCCGAGGT | CCTGGATCCG | TGTCTGTCTG | TGGTGCGCAG | CCTCCGTGCG | CTTCCGCTTA | 24360 |
| | CGGGGCCCGG | GGACCAAGGC | ACGAGTCCCA | GGGCTCTGAG | GGGCTCTGAG | GATCCTGGAC | CTTGGCCCAAC | 24430 |
| | GGCTCCTGCA | CCCCACCCCT | GTGGCTGCGG | TGGCTGCGGT | GACCCCGTCA | TCTGAGGAGA | GTGTGGGGTG | 24500 |
| 45 | AGGTGGACAG | AGGTGTGGCA | TGAGGATCCC | GTGTGCAACA | CACATGCGGC | CAGGAACCCG | TTTCAAACAG | 24570 |
| | GGTCTGAGGA | AGCTGGGAGG | GGTCTAGGTT | CCCGGGTCTG | GGTGGCTGGG | GACACTGGGG | AGGGGCTGCT | 24640 |
| | TCTCCCTTGG | GTCCCTATGG | TGGGGTGGGC | ACTTGGCCGG | ATCCACTTTC | CTGACTGTCT | CCCATGCTGT | 24710 |
| | CCCCGCCAGG | CCGAGCGTCT | CACCTCGAGG | GTGAAGGCAC | TGTTACAGCT | GCTCAACTAC | GAGCGGGCGC | 24780 |
| | GGCGCCCGGG | CCTCCTGGGC | GCCTCTGTGC | TGGGCTTGGG | CGATATCCAC | AGGGCTTGGC | GCACCTTCTG | 24850 |
| 50 | GCTGCGTGTG | CGGGCCAGG | ACCCGCCGCC | TGAGCTGTAC | TTTGTCAAGG | TGGGTGCCGG | GGACCCCGCT | 24920 |
| | GAGCAGCCCT | GCTGGACCTT | GGGAGTGGCT | GCCTGATTGG | CACCTCATGT | TGGGTGGAGG | AGGTACTCCT | 24990 |
| | GGGTGGGGCC | CAGGGAGTGC | AGGTGACCTT | GTCATGTTTG | AGGACACACC | TGGCACCTAG | GGTGGAGGCC | 25060 |
| | TTACAGCCTT | CCTGCAGCAC | ATGGGGCCGA | CTGTGCACCC | TGACTGCCCG | GGCTCCTATT | CCCAAGGAGG | 25130 |
| | GTCCCACTGG | ATTCCAGTTT | CCGTCAGAGA | AGGAACCCGA | ACGGCTCAGC | CACCAGGGCC | CGGTGCCTTG | 25200 |
| 55 | CACCCAGCTG | CTGAGCCAGG | GGTCTCCTGT | CCTGAGGCTC | AGAGAGGGGA | CACAGCCCGC | CCTGCCCTTG | 25270 |
| | GGGTCTGGAG | TGGTGGGGGT | CAGAGAGAGA | GTGGGGGACA | CCGCCAGGCC | AGGCCCTGAG | GGCAGAGGTG | 25340 |
| | ATGTCTGAGT | TTCTGCGTGG | CCACTGTGAG | TCTCCTCGCC | TCCACTCACA | CAGGTGGATG | TGACGGGGCG | 25410 |
| | GTACGACACC | ATCCCCCAGG | ACAGGCTCAC | GGAGGTCTAC | GCCAGCATCA | TCAAACCCCA | GAACACGTAC | 25480 |
| | TGCGTGCGTC | GGTATGCCGT | GGTCCAGAAG | GCCGCCCATG | GGCACGTCCG | CAAGGCCTTC | AAGAGCCACG | 25550 |
| 60 | TAAGGTTTAC | GTGTGATAGT | CGTGTCCAGG | ATGTGTGTCT | CTGGGATATG | AATGTGTCTA | GAATGCAGTC | 25620 |
| | GTGTCTGTGA | TGCGTTTCTG | TGGTGGAGGT | ACTTCCATGA | TTTACACATC | TGTGATATGC | GTGTGTGGCA | 25690 |
| | CGTGTGTGTC | GTGGTGCATG | TATCTGTGGC | GTGCATATTT | GTGGTGTGTG | TGTGTGTGGC | ACGTGTGTGT | 25760 |
| | CCATGGTGTG | TGTGCTGTGT | GTGTGCTATG | GTGTGTGTCT | GTGACACGTG | CATGTTTCATG | CTGTGTGCTG | 25830 |
| | GATGTCTGTG | ATGTGCTTAT | TTGTGGTGTG | TGTGCTGTATG | TGTCCGTGAC | ATATGCTGTG | CTATGGCATG | 25900 |
| 65 | GGTGTGTGTG | GCCCCCTTGG | CTTACTCCTT | CCTCCTCCAG | GCATGGTCCG | CACCATTTGT | CTCAGCTCTT | 25970 |
| | CGGGTGCTGG | TTTGGGGAGC | TCCACATTCA | GGGTCTCTAC | TTCTAGCATG | GGTGGCCCTG | TTCTGTCACT | 26040 |
| | GGGCTGGGGC | TTGGAGACTG | TAAGCCAGGT | TTGAGGATGAG | AGTAGGGATG | CTGGTGGTAC | TGCTTGGGAC | 26110 |
| | CCCTGGCACC | CCAGGACCCC | CAGTCTGGCT | TATGCCGGCT | CCATGAGATA | TAGGAAGGCT | GATTACAGGC | 26180 |
| | TGCGTCCCCG | GGACACACTC | CTCCAGAGGC | GGCCGGGGGC | CTTGGGGGCT | GGCAGGGGGT | AAAGGGGGCC | 26250 |
| | TGGGCTTGGG | TTCCACCCCA | GTGGTCTATG | GCACGCTGGA | GGGGTAAGCT | CTCAAAGTGC | TGCCAGGGCC | 26320 |
| 70 | GGGTGCAGAG | GTGAAGAAGT | ATCCCTGGAG | CTTCCGTCTG | GGGAGAGGCA | CATGTGGAAA | CCCACAAGGA | 26390 |
| | CCTCTTTCTC | TGACTTCTTG | AGCT | | | | | 26414 |

Contig 2:

| | | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|-------------|------|
| | TGTGGGATTG | GTTCATCATGT | GTGGGATAGG | TGGGGATCTG | TGGGATTGGT | TTTTATGAGT | GGGGTAACAC | 70 |
| | AGAGTTCAAG | GCGAGCTTTC | TTCTGTAGT | GGGTCTGCAG | GTGCTCCAAC | AGCTTTATTG | AGGAGACCAT | 140 |
| | ATCTTCCCTT | GAACATATGGT | CGGGTTTATA | GTAAAGTCAGG | GGTGTGGAGG | CCTCCCCTGG | GCTCCCTGTT | 210 |
| 5 | CTGTTTCTTC | CACTCTGGGG | TCGTGTGGTG | CCTGCTGTGG | TGTGTGGCCG | GTGGGCAGGG | CTTCCAGGCC | 280 |
| | TCCTTGTGTT | CATTGGCCTG | GATGTGGCCC | TGGCTACGCT | CCGTCCCTGG | AATCCCCCTG | CGAGTTGGAG | 350 |
| | GCTTCTCTTC | TTTCTTTTTT | TCTTCTTTTT | TTTTTTTTTT | TGATAACAGA | GTCTCGCTCT | TTTTTGCCCA | 420 |
| | GGCTGGAGTG | GTTTGGCGTG | ATCTTGGCTC | ACTGCAACCT | GTGCTTCTCT | AGTTCAAGCA | ATTCTCTTGC | 490 |
| | CTCAGCCTCC | CAAGTAGCTG | GAATTATAGG | CGCCACACAC | CATGCTGACT | AATTTTTGTA | ATTTTAGTAG | 560 |
| 10 | AGACGAGGTT | TCTCCATGTT | GGCCAGGCTG | GTCTCGAACT | CCTGACCTCA | GGTGATCCCT | CCACCTCGGC | 630 |
| | CTCCCAAAGT | GCTGGGATGA | CAGGTGTGAA | CCGCCGCGCC | CGGCCGAGAC | TCGCTTCTCT | CAGCTTCCGT | 700 |
| | GAGATCTGCA | GCGATAGCTG | CCTGCAGCCT | TGGTGCTGAC | AACCTCCGTT | TTCTTCTCTC | AGGTCTCGGT | 770 |
| | AGGGGTCTTT | CCATTTTCATG | ACTCTCTTCA | CAGAAGAGTT | TCAGTGTGTC | TGATTTCCCG | GCTGTTTCTT | 840 |
| | GCGTAATGGT | TGTCTGCTGT | TTATCGATGG | CCTCCTTTCA | TTTCTTTTAG | GCTTTGTTTA | TTGTTGTTTT | 910 |
| 15 | TCGGGTAGCT | TGAAGGAAAA | GTTTCGATTA | TGGATGTTTG | AACCTTCTTT | TCTAAACAAAG | CTGATGAAGT | 980 |
| | TGCCGTTTTT | CCTCTAAAGC | AGGGATCCCG | AGGCCCTGGG | CTGTGGAGTG | GCACCGGTCT | GGGGCTGTTT | 1050 |
| | AGGAACCCGG | GCGACAGCGG | GAGGCTAGGT | GGGGTGTGGG | GAGCCAGCGT | TCCCGCCTGA | GCCCCGCCCT | 1120 |
| | TCTCAGATCA | GCAGTGGCAT | GCGGTGCTCA | GAGCCGACCA | CACCTACTG | AGAAGTGTGC | GTGAGAGGGG | 1190 |
| 20 | TCTAGATTCT | GTGCTCCTTA | TGGGAATCTA | ATGCTGTATG | ATCTGAGGTG | GAACCGTTTG | CTCCCAAAAC | 1260 |
| | CATCCCCCTT | CCCACTGCTG | TCTGTGGGAA | AAATCGTCTT | CCACGAAACC | AGTCCCTGGT | ACCAACAATG | 1330 |
| | TTGGGGAAACC | TGTGCTAAAG | ACCTGCTTCA | GCAGCCTCTC | GTCAAGTGTG | ATATATTGGC | TTTTCTGTGT | 1400 |
| | TGAGTCCAGA | ATAATTACGG | ATTTCTGTGA | TGCTTTCCGC | CGACCTCAGA | CCCATGGGCT | ATTTGTGGGC | 1470 |
| | GTGTTGCTCT | CTCCTGGGTT | GGGAAGGGTG | CAGGCCCTAT | GTACCTTCTT | GTTACTGCCT | TCCAGGTTGG | 1540 |
| 25 | TTCTCAGGTT | TGAATCGTAC | TCGATGTGGT | TTTAGCCCTC | AGCCCTGCCC | CCAGCTCTCT | GGGGCTGAAG | 1610 |
| | AACATGCTGA | AGCACAGAGT | CACCGTGCGC | GTCTTTTGAT | GCCTCACAAG | CTCGAGGCCT | CCTGTGTCCG | 1680 |
| | TGTTAGTGTG | TGTCACGTGC | CTGCTCACAT | CCTGTCTTGG | GGACGACAGG | GCTTAGCAGG | TCCCGTAGTA | 1750 |
| | AATGACAAGT | GTCTGGGGG | AGTCTGCAGA | ATAGGAGTGT | GGGGTGCCCG | TCTCTCTCCC | GCGCTTCTAG | 1820 |
| | ACTCTTCTCC | TGCTGTGCTG | GTGGCTGCAC | CTGCATCCCT | GCAATCCCTC | CAGCACTGGG | CTGGAGAGGC | 1890 |
| | CCGGGAGCTG | GAGTGCCACT | TGTGCCACGT | GACTGTGGAT | GGCAGTCCGT | CACGGGGGCT | TGATGTGTGG | 1960 |
| 30 | TGACTGTGGA | TGGCGGTTTG | TCACAGGGGT | CTGATGTGTG | GTGACTGTGG | ATGGCGGTGC | TGGGGTCTGA | 2030 |
| | TGTGGTGAAT | GTGGATGGCG | GTGGTGGGGT | CTGATGTGTG | GTGACTGTGG | ATGGCGGTGC | TGGGGTCTGA | 2100 |
| | TGTGGTGAAT | GTGGATGGCG | GTGGTGGGGT | CTGATGTGTG | GACTGTGGAT | GGCGGTCTGT | GGGTCTGATG | 2170 |
| | TGTTGCTGAT | GGATGGCAGT | CGTGGGGTCT | GATGTGTGGT | GACTGTGGAT | GGCGGTCTGT | GGGTCTGATG | 2240 |
| 35 | TGTTGCTGAT | GTGGATGGCG | GTGGTGGGGT | CTGATGTGTG | GTGACTGTGG | ATGGCGGTGC | TGGGGTCTGA | 2310 |
| | TGTGTGGTGA | CTGTGGATGG | TGATCGGTCA | CAGGGGTCTG | ATGTGTGGTG | ACTGTGGATG | GCGGTCTGTT | 2380 |
| | GGTCTGATGT | GTGGTGAATG | TGGATGGTGA | TCGGTACAGC | GGGTCTGATG | TGTGGTGAAT | GTGGATGGCG | 2450 |
| | GTGCTGGGTT | GTGGTGAATG | TGGATGGTGA | TCGGTACAGC | GGGTCTGATG | TGTGGTGAAT | GTGGATGGCG | 2520 |
| 40 | GATGGCGATC | GGTCACAGGG | GTCTGATGTG | TGGTGAATGT | GGATGGCGGT | CGTGGGGTCT | GATGTGTGGT | 2590 |
| | GACTGTGGAT | GGCGGTCTGT | GGGTCTGATG | TGTGGTGAAT | GGATGGCGGT | CGTGGGGTCT | GATGTGTGGT | 2660 |
| | GACTGTGGAT | GGCGGTCTGT | GGGTCTGATG | TGTGGTGAAT | GGATGGCGGT | CGTGGGGTCT | GATGTGTGGT | 2730 |
| | GACTGTGGAT | GGCGGTCTGT | GGGTCTGATG | TGTGGTGAAT | GGATGGCGGT | CGTGGGGTCT | GATGTGTGGT | 2800 |
| 45 | GTGGTGAATG | TGGATGGCGG | GTGGTGGGGT | CTGATGTGTG | GTGACTGTGG | ATGGCGGTGC | TGGGGTCTGA | 2870 |
| | GTGGTGAATG | TGGATGGCGG | GTGGTGGGGT | CTGATGTGTG | GTGACTGTGG | ATGGCGGTGC | TGGGGTCTGA | 2940 |
| | GTGGTGAATG | TGGATGGCGG | GTGGTGGGGT | CTGATGTGTG | GTGACTGTGG | ATGGCGGTGC | TGGGGTCTGA | 3010 |
| | GTGGTGAATG | TGGATGGCGG | GTGGTGGGGT | CTGATGTGTG | GTGACTGTGG | ATGGCGGTGC | TGGGGTCTGA | 3080 |
| 50 | ATGTGTGGTG | ACTGTGGATG | GCGGTCTGTT | GGTCTGATGT | GGTGAATGTG | GATGGCGGTC | GTGGGGTCTG | 3150 |
| | ATGTGTGGTG | ACTGTGGATG | GCGGTCTGTT | GGTCTGATGT | GGTGAATGTG | GATGGCGGTC | GTGGGGTCTG | 3220 |
| | GGGTCTGATG | TGTGGTGAAT | GTGGATGGCG | GTGGTGGGGT | CTGATGTGGT | GACTGTGGAT | GGCGGTCTGT | 3290 |
| | GGGTCTGATG | TGTGGTGAAT | GTGGATGGCG | GTGGTGGGGT | CTGATGTGGT | GACTGTGGAT | GGCGGTCTGT | 3360 |
| | GTACACAGGG | TCTGATGTGT | GGTGAATGTG | GATGGCGGTC | GTGGTGGGGT | CTGATGTGGT | GACTGTGGAT | 3430 |
| 55 | GCGGTCTGAT | GGTCTGATGT | GTGGTGAATG | TGGATGGTGA | TCGGTACAGC | GGGTCTGATG | TGTGGTAGCT | 3500 |
| | TGGCGGTCTG | GGGTCTGATG | GTGGTGAATG | TGGATGGTGA | TCGGTACAGC | GGGTCTGATG | TGTGGTAGCT | 3570 |
| | GCAAGTGGAG | TCCCAAGGTG | GTCTGTAGCT | ACTTTGCGTC | CTCGGCCCTC | CGGCCCTCTC | TTCCCAAAAC | 3640 |
| | GAAGCTTCCC | AGGCGCTCTC | TGGGCTTTCAT | TGGGCTTTCAT | TGGGCTTTCAT | TGGGCTTTCAT | TGGGCTTTCAT | 3710 |
| 60 | GGAAGAAACA | AGTCCCCAGC | TCTGGCCGGG | GCAGGCCACA | TTTGTGGCTC | ATGCCCTCTC | CTCTGCCGGC | 3780 |
| | AGGTCTCTAC | CTTGACAGAC | CTCCAGCCGT | ACATGCGACA | GTTCGTGGCT | CACCTGCAGG | AGACCAGCCC | 3850 |
| | GCTGAGGAT | GCGGTCTGTC | TGGGCTGATG | GTGGGCTGTC | CCCTGCAGGG | TTGGGACAGG | ACTCCACGCA | 3920 |
| | GTGGGTCTCT | CCCTGGGCAA | TCACTGGGCT | CATGACCGGA | CAGACTGTTG | GCCCTGGGGG | GCACTGGGGG | 3990 |
| | GAATGAGCTG | TGATGGGGGC | ATGATGAGCT | GTGTGCCTTG | GCGAAATCTG | AGCTGGGGCA | TGCCAGGCTG | 4060 |
| 65 | CGACAGCTGC | TGCATTACAG | CACCTGTCTA | CGTTTGAAGT | CGCGGCTCTC | CTCCAGTTCC | GCACTGGGCT | 4130 |
| | TGTTTCATGAT | TTGCTAAATG | TCTTCTCTGC | CAGTTTTGAT | CTTGAGGCCA | AAGGAAAGGT | GTCCCCCTCC | 4200 |
| | TTTAGGAGGG | CAGGCCATGT | TTGAGCCGTG | TCCTGCCGAG | CTGGCCCCCT | AGTGTCTGGT | CTGAGGCCAA | 4270 |
| | AGGAACCTGT | TCCCTTCTCT | TAGGAGGACG | GGCGGTGTTT | GAGCCACGCC | CCGCTGAGCG | GGCTCTCTAG | 4340 |
| 70 | TGCTGGGTCT | GTCCACGTGG | CCCTGTGGCC | CTTTGCAGAT | GTGGTCTGTC | CACGTGGCCC | TGTGGCTCTT | 4410 |
| | TGCAGATGCC | TGTTAGCACT | TGCTCGGCTC | TAGGGGACAG | TCGTGTCCAC | CGCATGAGGC | TCAGAGACCT | 4480 |
| | CTGGGCGAAT | TTCTTGGGCT | CCAGGGGTGG | GGGTGGAGGT | GGCTTGGGCT | GCTGGGACCC | AGACCTCTGT | 4550 |
| | CCCGGCAGCT | GGGCAGCAAC | TCTTGGATCA | CATATGCCAT | CCGGGCCACG | GTGGGCTGTG | TGGGTGTGAG | 4620 |
| | CCGAGCTGGA | CCACAGAGTG | GCCAGAGGGA | GACGTTCTGT | GTCCACACCT | CTGCCAAGC | CCATGTGTGT | 4690 |
| 75 | CTGCAGAGAC | TCGGCCCGGC | CAGCCACGGA | TGCCCCCTCA | TTCCAGCCCA | GCCCCGCACT | TCATCAAAA | 4760 |
| | CACTGACCCC | AAAAGGGACG | GAGGGTCTTG | GCCACGTGGT | CCTGCCTGTC | TCAGCACCCA | CCGGCTCACT | 4830 |
| | CCCATGTGTC | TCCCGTCTGC | TTTCGCGAGG | CTCCTCCCTG | AATGAGGCCA | GCACTGGGCT | CTTCGACGTC | 4900 |
| | TTCTTACGCT | TCATGTGCCA | CCAGCCCGTG | GCATCAAGGG | GCAAGTGAGT | CAGGTGGGCA | GTTGCCATTG | 4970 |
| | CCCTGCGGGT | GGCTGGGCGG | GCTGGCAGGG | CTTCTGCTCA | CCTCTCTCCT | GCCCCCTCCC | CACTGNCCTT | 5040 |
| | CTGCCCGGGG | CCACAGAGTG | CTCCTTTTCT | GGCCCCCGGC | CCCTCCGGCT | CCTGGGCTGC | AGGCTCCCCG | 5110 |
| | GGCCCCGGA | ACATGGCTCG | GCTTGGCGGA | CCCGGAGCGG | AGCAGGTGCC | ACACGAGGCC | TGGAATGGC | 5180 |
| | AAGCGGGGTG | TGGAGTTGCT | CCTGCGTGGA | GGACGAGGGG | CGGGGGGTGT | GTCTGGGTCA | GGTGTGCGCC | 5250 |

| | | | | | | | | |
|----|------------|-------------|-------------|------------|-------------|-------------|------------|-------|
| | GAGCGTTTGA | GCCTGCAGCT | TGTCAGCTCC | AAGTTACTAC | TGACGCTGGA | CACCCGGCTC | TCACACGCTT | 5320 |
| | GTATCTCTCT | CTCCCGATAC | AAAAGGATTT | TATCCGATTC | TCATTCTGT | CCCTGTCTGT | TGACCCCGCT | 5390 |
| | GAGGGCGCGG | GCTCTTCTCT | CTGTGACTAG | ATTTCCCATC | TGGAAAGTGC | GGGGTTGACC | GTGTAGTTTG | 5460 |
| 5 | CTCCTCTCGG | GGGGCCTGTG | GTGGCCATGG | GGCAGGCGGC | CTGGGAGAGC | TGCCGTCACA | CAGCCACTGG | 5530 |
| | GTGAGCCACA | CTCACGGTGG | TAGAGCCACA | GTGCCTGGTG | CCACATCAGC | TCCTCTGGAT | TTTAAGTAAA | 5600 |
| | ACCACACACC | TCCCGGCAGG | CATCTGCCGT | CGACCTGTGT | TGTGCTGGGG | GAGAGTGGTA | GCACGGAGGA | 5670 |
| | AATTCGTGCA | CACTCAAGGT | CATCAGCAAG | GTCTCCGCA | GTGAGGTGGA | ACGTGGAGGC | CTCTCTCTGG | 5740 |
| | GATCGTCTCC | AGCGGATAAA | GGACTGTGCA | CAGCTTCGGA | AGCTTTTATT | TAAAAATATA | ACTATTAATT | 5810 |
| 10 | ATTGCATTAT | AAGTAATCAC | TAATGGTATC | AGCAATTATA | ATATTTATTA | AAGTATAATT | AGAAATATTA | 5880 |
| | AGTAGTACAC | ACGTCTCTGA | AAAACACAAA | TTGCACATGG | CAGCAGAGTG | AATTTTGGCC | GAGGGACACG | 5950 |
| | TGTGCACATG | TGTGTAAGCG | GCCCCCAGGC | CCACAGAATT | CGCTGACAAA | GTCACTCCCC | CAGAGAAGCC | 6020 |
| | ACCACGGGCC | TCCTTCGTGG | TCGTGAATTT | TATTAAGATG | GATCAAGTCA | CGTACCGTCC | ACGTGTGGCA | 6090 |
| | GGGCTTTGGG | GAATGTGAGG | TGATGACTGC | GTCTCATGTC | CCTGACAGAC | AGGAGGTGAC | TGTGTCTGTC | 6160 |
| 15 | CTGTCCCTAG | GACACGGACA | GGCCCCAAGC | TCTAGTCCCC | ATCGTGGTCC | AGTTTGGCCT | CTGAATAAAA | 6230 |
| | ACGCTCTCAC | AACTGTGTC | CCCCAAAACT | AAGAACAGAG | AGAGTTTCCC | ATCCCCATGT | CTCACAGGGA | 6300 |
| | CGTATCTGCT | TGCGTTGACT | CGCTGGGCTG | GCCGGACTCC | TAGAGTTGGT | GCGTGTGCTT | CTGTGCAAAA | 6370 |
| | AGTGCAGTCC | TCTTGCCCAT | CACGTGTGATA | TCTGCACAGC | CAAGGAAAGC | CTCTTTTCTT | TTCTTTCTTT | 6440 |
| 20 | TTTTTTTTTT | GAGACGGAAC | GTCACTGTTG | TCTGCCTGGG | CTTGAGTGCA | GTGGCGGATG | CTCACAGGGA | 6510 |
| | TGCAACCTCC | GCCTCCCGGG | TTCCAGCATT | TCTCCTGCCT | CAGCTCCCG | AGCAGCTGAG | ATTACAGGCA | 6580 |
| | CCCACCCCTC | GCGCCTGGCT | AATTTTTGTA | TTTTTAGTAG | AGAGGGGTTT | TTGCCATGTT | GGCCAGGCTG | 6650 |
| | GTCTCGAACT | CCTGACCTCA | GGTGATCCAC | CCACCTCGGC | CTCCCAAAGT | GCTGGGATTA | ACGTGTGGCA | 6720 |
| | CCATCACGCC | CAGCCGGAAA | GCCTCTTTTT | AAGGTGACCA | CCTATAGCGC | TTCCCGAAAA | TAACAGGTCT | 6790 |
| 25 | TGTTTTTGCA | GTAGGCTGCA | AGCGTCTCTT | AGCAACAGGA | GTGGCGTCTC | GTGGGCTCTG | GGGATGGCTG | 6860 |
| | AGGGTCGCGT | GGCAGCCATG | CCTTCTGTGT | GCATCTTTAG | GTTCCACGGG | GTTCTACTGC | TCTCACTGTT | 6930 |
| | TGTCTGAAAA | CGCACCCCTT | GCATCCTTGT | TTGGAGAGTT | TCTGCTTCTC | GTTGGTCATG | CTGAAACTAG | 7000 |
| | GGGCAAGGTT | GTATCCGTTG | GCGCGCAGCG | GCTACATGTA | GGGTTCATGAG | TCTTTCACCG | TGGACAAATT | 7070 |
| 30 | CCTTGAAAAA | AAAAAAAGAA | GTCCGGTTAA | GCATTACATC | CGGGTCAAGT | GTCTGGTTCT | GTGAATAAAC | 7140 |
| | TCTAAGATTT | AAGAAACCTT | AATGAAAGAA | AACCTTGATG | ATTGAGAGCA | AGGATGTGGT | CACACCTGTG | 7210 |
| | GCTGGATCTG | TTTCAGCGCC | CCCAGTGCAT | GGTGAGAGTG | GGGAGCAGGG | ATTGTTTGGT | CAGAGGTCTC | 7280 |
| | ATCTGGTATG | TTTCTGAGGT | GTTTGCCGGC | TGAATGGTAG | ACGTGTCGTT | TGTGTGTATG | AGGTTCTGTG | 7350 |
| | TCTGTGTGTG | GCTCGGTTTG | AGTGTACGCA | TGTCCAGCAC | ATGCCCTGCC | CGTCTCTCAC | CTGTGTCTTC | 7420 |
| 35 | CCGCCCCAGG | TCCTACGTCC | AGTGCCAGGG | GATCCCCGAG | GGCTCCATCC | TCTCCACGCT | GCTCTGCAGC | 7490 |
| | CTGTGCTACG | GCGACATGGA | GAACAAGCTG | TTTGCGGGAG | TTCCGGCGGA | CGGGTGAAGC | CTCTCTTCCC | 7560 |
| | CCAGGGGGGC | TTGGGTGGGG | GTTGATTTGC | TTTTGATGCA | TTCACTGTTA | ATATTCCTGG | TGCTCTGGAG | 7630 |
| | ACCATGACAG | CTCTGTCTTG | AGGAACCCAG | CAAGCTTGCA | GCCCCTTCTT | GGTATGAAGC | CGCACGGGAG | 7700 |
| | GGGTTGCATG | CCCTGAGGAG | TGCCGGCTCC | ACCGAGGCTC | CCATGTCAGG | TTGTCACAGC | AGGCCTCAGG | 7770 |
| | GCTCAGCAGG | CGGGAGGGCC | GCTGCCCTGC | ATGATGAGCA | TGTGAATTCA | ACACCGAGGA | AGCACACCAG | 7840 |
| 40 | CTTCTGTCTC | GTCAACCCAG | TTCCGTTAGG | GTCTTTGGGG | AGATGGGGCT | GGTGCCAGCT | GAGGCCCCAC | 7910 |
| | ATCTCCGAGC | AGGCCCTCGA | CAGGTGGGCT | GGACTGGGGC | CCTCTTACAG | CCATTGCCCA | TCCCTCTTGC | 7980 |
| | ATGGGGTCTA | CACCCAAGGA | CGCACACACC | TAAATATCGT | GCCAACCTAA | TGTGGTTCAA | CTCAGCTGGC | 8050 |
| | TTTTATTGAC | AGCAGTTACT | TTTTTTTTTT | TAACTACTTA | AGTTCTAGGG | TACATGTGCA | AGCAGTGCAG | 8120 |
| | GTTAGTTACA | TATGTATACA | TGTGCCATGT | TGGTGTGCTG | CACCCATTAA | CTCATCATTT | CGAATTAGGT | 8190 |
| 45 | TATCTCTCTA | TGCTATCCCT | CCCCACTCCC | CCCATCCCAT | GACAGGCCCT | GGTGTGTGAT | GTTCCCCACC | 8260 |
| | CTGTGTCCAA | GTGTTCTCAT | TGTTTCAGTT | CCACCTGTGA | GTGAGAACAT | GTGGTGTGTT | GTTTTCTTTC | 8330 |
| | CTTGCAATAG | TTTGCTCAGA | GTGATGGTTT | CCAGCTTCCG | CCATGTCCCT | ACAAAGGACA | TGGAATCATC | 8400 |
| | CTTTTTTATG | ACTGCATAGT | ATTCGGTGGT | GTATATGTGC | CACATTTTCT | TAATCCAGTC | TATCATCGAT | 8470 |
| 50 | GGACATTTGG | GTTGGTTGCA | AGTCTTTGCT | ACTGTGAATA | GTGCCGCAAT | AAACATACGT | GTGCATGTGT | 8540 |
| | TTTATATAGA | GCATGATTTA | TAACTCTTTG | GGTATATACC | CAGTAATGGG | ATGGCTGGGT | TCCCTAAGAG | 8610 |
| | TTCTAGTTCT | AGATCCTTGA | GGAATCACCA | CACGTCTCTC | CACAATGGTT | GAAC TAGTTT | ACACTCCCAC | 8680 |
| | CAACAGTGTA | AAAGTGTCTT | GGTGCTGGAG | AGGATGTGGA | CAGCAGTTAT | TTTTTTATGA | AAATAGTATC | 8750 |
| | ACTGACAGTA | CAGACAGTTA | GTGAAGGATG | CCTGCAGGAG | ACACAGCCAT | TTCTCTCGAA | TTCTCTCGAA | 8820 |
| | GACTCCGGGT | TTTTCTCTGT | CATCTTTTGA | AACCTTAGCT | CCAATTATAG | CATGTACAGT | GGATCAAGGT | 8890 |
| 55 | TCTTCTTCTA | TAAGGTTCAA | GTTCTAGATT | GAAATAAGTT | TATGTAACAG | AAACAAAAAT | TTCTTGTACA | 8960 |
| | CACAACCTGC | TCTGGGATTT | TGAGGAAAGT | GTCTCTGAGC | TGGCGGCACA | CTGGTCAAGC | CTGCTGGACA | 9030 |
| | GGATACCTCT | GGCCCATGGT | CATGGGGCGC | TGGGCTTGGG | CCTGAGGGTC | ACACAGTGCA | CCATGCCCCG | 9100 |
| | CTTCTGTGGG | ATAGGATCTG | GGTCTCGGAT | CATGCTGAGG | ACCACAGCTG | CCATGCTGGT | AAAGGGCACC | 9170 |
| | ACGTGTGCTA | GAGGGGGCGA | GGTTCCACAG | CCCAGCTTTT | TTACCGTCTT | CAGTTATTTT | TCCCTAAGAG | 9240 |
| 60 | TCTGAGAAGT | GGGGCCGCGC | CTGATGGCCT | TCGTTCTGCT | TCAGCTGGCA | CAGAATTGCA | CAAGCTGATG | 9310 |
| | GTAACACATG | AGTACTTATA | ATGAATGAGG | AATTGCTGTA | GCAGTTAACT | GTAAGAGAGT | CGTCTGTTGG | 9380 |
| | AAAGAAATTT | AAGTTTTTCA | TTTTAACCGCT | TTGGAGAATG | TTACTTTTAT | TATGGCTGTG | TAAATTTGTT | 9450 |
| | GACATTCAGT | CCCTCGTAGA | CAGATACTAC | GTAAAAGTGT | TAAAGTTAAC | CTTGCTGTGT | ATTTTCCCTT | 9520 |
| | ATTTTAGGCT | GCTCCTGCGT | TTGGTGGATG | ATTTCTTTGT | GGTGACACCT | CACCTCACCC | ACGCGAAAAA | 9590 |
| 65 | CTTCTCTCAG | TGAGGCCCGT | GCCGTGTGTC | TGTGGGGACC | TCCACAGCCT | GTGGGCTTTG | GATGTGAGCC | 9660 |
| | CCCGGTGTCC | TGCCCTTGGC | ACCGCAGCGT | TGTCTCTGCC | AAGTCTCTCT | TCTCTGCCCG | TGCTGGATCC | 9730 |
| | GCAAGAGCAG | AGGCGCTTGG | CCGTGACACC | AGGCTTGGGG | GCGCAGGGGC | ACCTTCGGGA | GGGAGTGGGT | 9800 |
| | ACCGTGCAGG | CCCTGGTCTT | GCAGAGACGC | ACCCAGGTTA | CACAGTGGGT | GAGTGCAGGC | GAGTGACCTG | 9870 |
| | CTCCTGCTGC | TCTTTGGAAG | GTCAAGAGTG | GCGGCTCCTG | GGGCCCCAGT | GAGACCCCCA | GGAGCTGTGC | 9940 |
| 70 | ACAGGGCCTG | CAGGGCCGAG | GCGGCAGCCT | CCTCCCCAGG | GTGCACCTGA | GCCTGCGGAG | AGCAGGAGCT | 10010 |
| | GCTGAGTGAG | CTGGCCACAC | GCGTTCGCTG | CGGTACGTTT | CCTGCGTGGG | GTTGTTTGGG | ATCCGTGGGA | 10080 |
| | GAATTTGGAT | TTGCTGAGTG | CTGCTGTCTT | GAAACACGGA | GATGGCTAGG | AGTGGGTTTC | AGAGTTGATT | 10150 |
| | TTTGTAATTC | AAACTAAAT | CAGGCACAGG | GGAGCTGGCC | TCAGCACAGG | GGATTGTCCA | ATGTGTTCCC | 10220 |
| 75 | CCTCAAGGGC | GCCCCACAGA | GCCGTTGCGC | TTGTTTTAAA | GTGCGATTTC | ACGAGGGACG | AGAAACCTTG | 10290 |
| | AAAGCTGTAA | AGGGAACCCCT | CAGAAAATGT | GGCCGCCAGG | GGTGGTTTCA | GGTGCTTTGC | TGGGCTGTGT | 10360 |
| | TTTGTAATTC | CCATTTGGAC | CCGCCCTCCA | AGTCCACCTC | CCAGGTCCAC | CCTCCAGGGC | CGCCCTGGGG | 10430 |
| | TGGGGGTATG | CCTGGCGTTT | CTTGTGCGGC | AGCCCGGAGC | ACAGCAGGCT | GTGCACATTT | AAATCCACTA | 10500 |
| | AGATTCACTC | GGGGGGAGCC | CAGGTCCCAA | GCAACTGAGG | GCTCAGGAGT | CCTGAGGCTG | CTGAGGGGAC | 10570 |
| | AGAGCAGACG | GGGAACGCTG | CTTCTGTGTG | GCAAGTTCTT | GAGGCTGTGT | GCCAGGGAGG | TGGCTCAGAG | 10640 |
| | TGTATGTTTG | GGTCCCACCG | GGGGCAGAAC | TCTGTCTCTG | ATGAGTCCGC | AGCCATGTAA | CAGGAAGGGG | 10710 |

002200 9422550

| | | | | | | | | |
|----|------------|------------|------------|------------|-------------|------------|-------------|-------|
| | TGGCCACAGG | GAGCTGGGAA | TGCACCAGGG | GAGCTGCGCA | GCTGGCCGAG | GTCCCAGGGC | CAGGCCACAG | 10780 |
| | GAAGGGCAGG | GGGACGCCCG | GGGCCACAGC | AGAGGCCCGA | GGAAGGGAAG | GGGATGCCCA | GGCCAGAGCA | 10850 |
| | GAGGCTACCG | GGCACAGGGG | GGCTCCCTGA | GCTGGGTGAG | CGAGGCTCAT | GACTCGGCGA | GGGAACCTCC | 10920 |
| 5 | TTGACGTGAA | GCTGACGACT | GGTGTGCCCC | AGCTCACAGC | CCAGCCAGGT | CCCGCGCCTG | AGCAGGAACT | 10990 |
| | CAGAACCCTC | CCCTTTGTCT | AAAGCACAGC | AGATGCCTTC | AGGGCATCTA | GGAGAAAACA | GGCAAAGTCG | 11060 |
| | TTGAGAAACG | TCTTAAAGA | AGGTGGGATG | GTGGCAATTT | CTTGTCAGAG | TTTGTAGCTG | CCCCGGACCA | 11130 |
| | CAGATGAGTC | TATAACGGGA | TTGTGGTGT | GCCATGGGGA | CACATGAGAT | GGACCATCAC | AGAGGGCCACT | 11200 |
| | GGGGCTGCAC | CTCCCATCTG | AGTCCTGGCT | GTCCCGGGTC | CAGGCCAGGT | TCTTGATATC | TCACCTACCT | 11270 |
| 10 | GTCCTGCCCG | GGAGACAGGG | AAAGCACCCC | GAAGTCTGGA | GCAGGGCTGG | GTCCAGGCTC | CTCAGAGCTC | 11340 |
| | CTGCCAGGCC | CAGCACCTCG | CTCCAAATCA | CCACTTCTCT | GGGGTTTTC | AAAGCATTTA | ACAAGGGTGT | 11410 |
| | CAGGTTACCT | CCTGGGTGAC | GGCCCCGCAT | CCTGGGGCTG | ACATTGCCCC | TCTGCTTAG | GACCCTGGTC | 11480 |
| | CGAGGTGTCC | CTGAGTATGG | CTGCGTGGTG | AACCTGCGGA | AGACAGTGGT | GAACCTCCCT | GTAGAAGACG | 11550 |
| | AGGCCCTGGG | TGGCACGGCT | TTTGTTCAGA | TGCCGGCCCA | CGGCCCTATC | CCCTGCTGCG | GCCTGCTGCT | 11620 |
| 15 | GGATACCCCG | ACCCTGGAGG | TGCAGAGCGA | CTACTCCAGG | TGAGCGCAC | TGGCCGGAAG | TGGAGCCCTG | 11690 |
| | SCCCGGCTGG | GGCAGGTGCT | GCTGCAGGGT | ACCTCTGCTT | CCGTGCTGGG | CCGTGCTGGG | CAGCGGACCA | 11760 |
| | CCAATCCCAA | AGGGTCAGAG | GCCACAGGGT | GCCCCCTGTC | CCATCTGGGG | CTGAGCAGAA | ATGCATCTTT | 11830 |
| | CTGTGGGAGT | GAGGGTGCTC | ACAACGGGAG | CAGTTTTCTG | TGCTATTTTG | GTAAGGGAAG | ATGGTGCCAC | 11900 |
| 20 | AGACCTGGGT | GCACTGAGGT | GTCTTCAGAA | AGCAGCTGGA | ATCCGAACCC | AAGACGCCCG | GGCCCTGCTG | 11970 |
| | GGCGTGAGTC | TCTCAAACCC | GAACACAGGG | GCCCTGCTGG | GCATGAGTCC | CTCTGAACCC | GAGACCCTGG | 12040 |
| | GGCCCTGTCT | GGCGTGAGTC | TCTCCGAACC | CAGAGACTTC | AGGGCCCTTT | TGGGCGTGAG | TCTCTCCGCT | 12110 |
| | GTGAGGCCCA | CACCTCAAGG | CTCATCCACA | GTCTACAGAT | TGCCATGAGT | TCATGATCAC | GTGTGAGCCA | 12180 |
| | TCAGGGGACA | GGGCCATGGT | GTGGGGGGGG | TCTCTACAAA | ATTCTGGGGT | CTTGTTCCTC | CAGAGCCCGA | 12250 |
| | GAGCTCAAGG | CCCCGTCTCA | GGCTCAGACA | CAAAATGAAT | GAAGATGGAC | ACAGATGCAG | AAATCTGTGC | 12320 |
| 25 | TGTTTTCTTT | ATGAATAAAA | AGTATCAACA | TTCAGGCGAG | GGCAAGGTGG | CTCACACCTA | TAATCCAGC | 12390 |
| | ACTTTGGGAG | GCCGAGGTGG | GTGGATCACT | TGAGGCCAGG | AGTTTGAGGC | CAACCTAACC | AACATAGTGA | 12460 |
| | AATTTCAATT | CTACTTAAAA | AATACAAAAA | TTAGCCTGGC | CTGGTGGCAC | ACGCCTGTAG | TCCCCGCTAT | 12530 |
| | GCGGGAGGCT | GAGGCAGGAG | AATCATTTGA | ACCCAGGAGG | CAGAGGTTGC | AGTGAGCCGA | GATCACACCA | 12600 |
| | CTGCACTCCA | GCCTGGGCAA | CAGAGTGAGA | CTTCATCTTA | AAAAAAGTATC | AGCATTCCAA | AGCATTCCAA | 12670 |
| 30 | AACCATAGTG | GACAGGTGTT | TTTTTATTCT | GTCTTCGAT | AATATTACT | GGTGTGTGTC | TAGAGGCCGG | 12740 |
| | AACCTGGGGT | GCCTTCTCTC | GAAAGGCACA | CCTTCATGGG | AAGAGAATA | AGTGGTGAAT | GGTGTGAAT | 12810 |
| | CCAGAGGTTT | AAACTGGGGT | CCTGTCGTTT | TGAGTTAACA | GTCCAGATCT | GGACTTTGCC | TCTTTCCAGA | 12880 |
| | ATGCTCCCTG | GGGTTTGGCT | CATGGGGGAG | CAGCAGGTGT | GGACACCCCT | GTGATGGGGG | AGCAGCAGGT | 12950 |
| 35 | GCAGACGCC | TCATGATGGG | GGAGTGGCAG | GTGCCAGCAC | CCTTGTGTC | GGTGGCCAGC | ATGTCCTGT | 13020 |
| | TGCAGCTCCC | TCCCCACAAG | GATGCCGGTC | TCCTGTGCTC | CCCACAGTCC | CTGCTTCCCT | CTCACAGCCT | 13090 |
| | TACCTGGTCT | TGGCCTCCAC | TGGCTTTGTC | TGCATGATTT | CCACATTTCC | TGGGCTCCCA | GCACCTCTTC | 13160 |
| | GCCTCTCCCA | TGCACCTCTG | CAGTGTCTGC | CATACCAATC | AGCTGTGAAC | TGTCCACTGC | TTATTTGTCT | 13230 |
| | CCCCATGAAA | TGTATTTTTT | AGGACAGGCA | CCCCTGGTTC | CAGCCTCTGG | CACAGCATCA | GTGAATGTTA | 13300 |
| | TTGAAGGACA | AAGGACAGAC | AAACAAATCA | GGAAATGGG | TTCTCTCTAA | ACACATTGCA | AGGCCACAGA | 13370 |
| 40 | GGCTAGTGCA | GGATGGGTGG | GCATCAGGTC | ATCAGATGTG | GGTCCAATGC | CAGAATATT | TGTGCTCCCA | 13440 |
| | AAGGCCACTT | GGTCAGAGTG | TGTGCTTGCA | GAGGTGGCTC | TAAAGCTCA | GCAGTGAGAG | CAGTGGTTCT | 13510 |
| | CCATCTCAG | GGTGAACTCA | CATCCTCTGT | GTCTGAAAGT | TACAGCAGAG | GCTTGAAGGG | CATCTGGGAG | 13580 |
| | AAGAAAACAG | GCAAAATGAT | TAAAGAAAGT | GAAAGAGGAA | AAGTGGAAG | ATGGGAATTT | TCTTGTCCAG | 13650 |
| | ATTTTACTCT | CCCAAACAC | AGCTCAGATG | GTAGAATGTG | GTCAGAACTG | ATGGACAGAA | CAATAGAAC | 13720 |
| 45 | AAACGGGAAG | CCTATCTCTC | AGAAACGTGT | GTAAATGGG | TATGTGGCAC | AGCTGATGGA | AAAGAGAGTG | 13790 |
| | TGTGTGTAAT | TTTTTTTCT | GAGAAAACCT | ACTGGAAGCA | AATAAGTTGT | GTCTTTTACA | CATATACCAG | 13860 |
| | AGCAGATTCT | AGGTAGAAGA | GGAGACACAT | GCAACAACA | CCAGCAACAG | AAATAAAACA | AAAGACTCAA | 13930 |
| | AGGGAAGGGA | GGTGAACGTT | CCCTGGTTTG | GTGTTGGGGA | AGGACACACA | GGGAGGCGGA | TGAAACCACT | 14000 |
| | GAGGCAACCG | GCATTGCTTT | CACCTGCAGG | AAACTGCAGT | TGCCTGAGCC | ACAGTGAAAA | TGGCCATTCC | 14070 |
| 50 | CTGGAGCGTT | TGTGCACGTG | ATTTATTTAA | GGCGCCCTGT | GAGGTCTGTC | ACATTCTATC | TCTCACTTTG | 14140 |
| | TTCTCTTAAC | CACCTGAGAG | GTAGAGGAGG | AAAGGCTCCA | GGGGAGCAGC | CGCCCTTGGT | CACCCAGCTG | 14210 |
| | GCAAAAGGCA | TGCATGATTG | CAGCCTGGCC | TCTGTCTCCG | GGGCGCTTGC | TCTGCCCGAG | GGCCACACAC | 14280 |
| | AAGTCAGACC | CATAGGCTCA | GGGTGAGCCG | GAGCCCAAGG | TCGTGTGGG | GATGGCTGTG | AAAGAAGAAA | 14350 |
| | TGGACGTCTG | ATGCACACTT | GGGAAGGTCC | TACCAGCAGC | GTCAAAGAAA | TCGATGTGAA | ACTGACAGCG | 14420 |
| 55 | AGACCCATCC | CTCAAAGAAA | CGCACGTGAA | ACTGATGGCG | AGACCTGTCC | CCATCCCTCA | TGCTGGCTCC | 14490 |
| | TTTTCTGGGC | TTGCCAAGAG | CCAGCATCAG | GTTGAGGCAA | GCTGGAAGAA | CTTTTCTGGA | AAGCAGCTTG | 14560 |
| | TTTGATGTTA | AGTCCTCACA | ATGTCCTGTG | TCTTCCAGT | AATTCACCTT | CTGAAGTGAC | CAGACATTAT | 14630 |
| | CACGGGTCTT | ATTTACCATT | TCCAGTGTTT | CAGGCAGGGG | GACTTGCCAC | AGCAAGTCAC | GAACCTGCC | 14700 |
| | AAATACAGGG | CTAAGGAGAT | ATTATGCATC | ACAAAACCTT | CTCTGCCATT | AAACATTTTT | CAAAGAATTT | 14770 |
| 60 | TTGAAGAATG | TTTAATGGCA | CAAAACGTTT | ATTTCAATGT | AGCAGTGTTT | AAAGCTGGAT | GTAAAAGAAC | 14840 |
| | ACACCCAGAG | AGCCTGCCGT | GAATGTCTAT | TGTGTTTATC | TTTGGACATG | GACATACATG | GGCAGTGAGT | 14910 |
| | GGTGGTGAGG | CCCTGGAGGA | CATCGGTGGG | ATGCCTCCAT | CCTGCCCTTC | TGGAGACACC | ATGTGTGCCA | 14980 |
| | CGTGCACTCA | CTGGAGCCCT | GTTTAGCTGG | TGCCACCTGG | CTCTTCCATC | CCTGAGATT | AAACACAGTG | 15050 |
| 65 | AGATTCCTCA | CGCCCAACTC | AGTGTCTCTC | CACAAAACAA | CTGAGTCACA | CCTGTGTTCA | CTCGAGGGAG | 15120 |
| | GCCCGGGAGC | CAGGGCTCCA | CAGTTTATTA | TGTGTTTTTG | GCTGAGTTAT | GTGCAGATCT | CATCAGGGCA | 15190 |
| | GATGATGAGT | GCACAAACAC | GGCCGTGCGA | GGTTTGGATA | CACCTCAACAT | CACCTGAGCC | GTCTGCTGCT | 15260 |
| | AGTTTGGTCA | TGCAGAGTCT | GGATGGCATG | TAGCATTTGG | AGTCATGGA | GTGAGACACC | AGCCCCCTCG | 15330 |
| | GGCTGCAGCG | CATGCCCCAG | GCAGGACAAG | GAAGCGGGAG | GAAGGCAGGA | GGCTCTTTGG | AGCAAGCTTT | 15400 |
| | GCAGGAGGGG | GCTGGGTGTG | GGGCAGGCAC | CTGTGTCTGA | CATTCCCCCC | TGTGTCTCAG | CTATGCCCGG | 15470 |
| 70 | ACCTCCATCA | GAGCCAGTCT | CACCTTCAAC | CGCGGCTTCA | AGGCTGGGAG | GAACATGCGT | CGCAAACTCT | 15540 |
| | TTGGGGTCTT | GCGGCTGAAG | TGTCACAGCC | TGTTTCTGGA | TTTGCAGGTG | AGCAGGCTGA | TGGTCAGCAC | 15610 |
| | AGAGTTCAGA | GTTCAGGAGG | TGTGTGCGCA | AGTATGTGTG | TGTGTGTGTG | CGCGCTGTCC | TGCAAGGCTG | 15680 |
| | ATGTTGACTG | GCTGCACGTA | AGAGTGCACA | TGTACGCGTA | TACACGTGAG | CACATACATG | TGTGATGTG | 15750 |
| | TGTACATGAA | GGCATGGCAG | TGTGTGCACA | GGTGTGCAAG | GGCACAAGTG | TGTGCACATG | CGAATGCACA | 15820 |
| 75 | CCTGACATGC | ATGTGTGTTC | GTGCACAGTC | GTGTGGGCGT | TCACGTGAGG | TGCATGCGTG | TGGGTGTGCA | 15890 |
| | GTGTGAGTAG | CATGTGTGCA | CATAACATGT | ATTGAGGGGT | CCTCGTGTTC | ACCCCGCTAG | GCTCTCAGCA | 15960 |
| | CCAGTGCCAC | TCCTTACAGG | ATGAGACGGG | GTCCAGGCC | TTGGTGGGCT | GAGGCTCTGA | AGCTGCAGCC | 16030 |
| | CTGAGGGCAT | TGTCCCATCT | GGGCATCCGC | GTCCACTCCC | TCTCTGTGG | GCTTCTGTGT | CCACTCCCCC | 16100 |
| | TCTCTGTGG | GCATTTACAT | CCACTCCACT | CCCTCTCTCC | TGTGGGCATC | CGCGTCCACT | CCCCCTCTCT | 16170 |

| | | | | | | | | |
|----|-------------|-------------|-------------|------------|-------------|-------------|-------------|-------|
| | GTGGGCATCT | GCGTCCACCT | CCCCCTCTCTG | TGGGCATTGG | CGTCCACTCC | CTCTCTCTGGT | TCCTTCTCTGT | 16240 |
| | CTTGCGCCGAG | CCTCGGGGGC | AGGCAGATGA | CACAGAGTCT | TGACTCGCCC | AGGGTGGTTT | GCAGCTGCGG | 16310 |
| | GGTGAGGGCC | AGGCCGGAT | TCACTGGGA | GAGGGATAGT | TTCTTGTCAA | AATGTTCTCT | TTTCTTGTTC | 16380 |
| 5 | CATCTGAATG | GATGATAAAG | AAAAAAGTAA | AAACTTAAAA | TCCCAGTAGC | GTTTCTTACCG | TTTCTCACTG | 16450 |
| | TTTCTTGCGG | ACTCTAGGTT | AACAGCCTCC | AGACGGTGTG | CACCAACATC | TACAAGATCC | TCCTGCTGCA | 16520 |
| | GGCGTACAGG | TGAGCCGCCA | CCAAGGGGTG | CAGGCCAGC | CTCCAGGGAC | CTCTCCGCGT | CTGCTCACCT | 16590 |
| | CTTGACCCGG | GCTTCACTTT | GGAACTCTGT | GGTTTATAGG | CGCAAGGAATG | TCTTACGTTT | TCAGTGGTGC | 16660 |
| | TGCTGCCTGT | GCACAGTTCT | GTTTCGCGTGG | CTCTGTGCAA | AGCACTGTGT | CTCCATCTCT | GGGTAGTGGT | 16730 |
| 10 | AGGAGCCGGT | GTGGCCCCAG | GTGTCCCCAC | TGTGCTCTGT | CAGTGGCCGT | GGGACGTCTAT | GGAGGCCATC | 16800 |
| | CCAGGGCAGC | AGGGGCATGG | GGTAAAGAGA | TGTTTATGGG | GAGTCTTAGC | AGAGAGGAGT | GGGAAGGTGT | 16870 |
| | CTGAACAGTA | GATGGGAGAT | CAGATGCCCG | GAGGATTGGG | GGTCTCAGCA | AAGAGGGCCG | AGGTGGGTGC | 16940 |
| | AGGTGAGGGT | CGCTGGCCCC | ACCCCGGGGA | AGGTGCAGCA | GAGCTGTGGC | TCCCCACACA | GCCCCGCCAG | 17010 |
| | CACCTGTGCT | CTGGGCATGG | CTGTGCTCCT | GGAACGTTTC | CTGTCTGGC | TGGTCAGGGG | GTGCCCTGTC | 17080 |
| 15 | CAAGAATCGA | CAACTTTATC | ACAGAGGGAA | GGGCCAATCT | GTGGAGGCCA | CAGGGCCAGC | TTCTGCCTGG | 17150 |
| | AGTCAGGGCA | GGTGGTGCCA | CAAGCCTCGG | GGCTGTACCA | CCAGGCGAGT | GGGACCCACA | GGCCCGGGCC | 17220 |
| | TCCACCTCAA | CAGCGCTCCC | GAGCCACTGG | GAGCTGAATG | CAAGGAGGCC | GAAGCCCTCG | CCCCATGAGG | 17290 |
| | GCTGAGAAAG | AGTGTGAGCA | TTTGTGTTAC | CCAGGCGCCA | GGCTGCAGCA | ATTACCGTGC | ACACTTGATG | 17360 |
| | TGAAATGAGG | TCGTCTGCTA | TCGTGGAAAC | CCAGCAAGGG | CTCACGGGAG | AGTTTTCATC | TACAAGGTGC | 17430 |
| 20 | TACCATGAAA | ATGTTTTTTA | ACCCGAGTGC | TTTGGCCTTC | ATGCTTGGC | AGGGAGGGCA | GAGCCACAGC | 17500 |
| | TGCATGTTAT | CGCCTTTGCA | CCAGCTCCAG | AGGCTTGGGA | CCAGGCTGTC | TCAGTTCAGC | GGTGCCTCCG | 17570 |
| | GCTCAGACCG | CCCTCTCTCT | TGCTTCTCTT | CTCTGCCTCA | AACTTTCCTT | CGTTTGCATC | TCCCTGACGC | 17640 |
| | GTGCTTGGGC | CTCTGTGCAA | GCTGCTTGAC | TCCTTTCGGG | AAACCCCTGG | GTGCTGCTGG | ATACAGGTGC | 17710 |
| | CACCTGAGAG | TGAGAGTGT | TGACACTGTG | GTTGACCCCA | GGGTCCAGCT | GGCGTGCTGT | GGGCCTCCTT | 17780 |
| 25 | GGGCCATGAT | GAGGTACAGT | GAGTTTTCCC | AGGTGAAAC | TCCTGGGAAA | CTCCCAGGGC | CATGTGACCT | 17850 |
| | GCCACCTGCT | GCTCCCATAT | TCAGCTCAGT | CTTGTCTCTA | TTTTCCACC | AGGGTCTCTA | GCTCCGAGGA | 17920 |
| | GCTCCCGTAG | AGGGCCTGGG | CTCAGGGCAG | GGCGGCTGAG | TTTTCCACC | CATGTGGGGA | CCCTTGGGTA | 17990 |
| | GTCGTTGAT | TGGGTAGCCC | TGAGGAGGCC | GAGATGCGAT | GGGCCACGGG | CCGTTTCCAA | ACACAGAGTC | 18060 |
| | AGGCAGCTGG | AAGGCCCAGG | AATCCCCCTC | CCTCAGGGCA | GGAGTGGGAG | AACGGAGAGC | TGGGCCCGGA | 18130 |
| 30 | TTTCACGGCA | GCCAGGCTGC | AGTGGGCGAG | GCTGTGGTGG | TCCACGTGGC | GCTGGGGGCG | GGGTCTGATT | 18200 |
| | CAAAATCCGT | GGGGCTCGGC | CTTCTCTGGC | CGTGTGGCC | CGCCCTCCAC | ACGGGCTTGG | GGTGGACGCC | 18270 |
| | CCGACCTCTA | GCAGGTGGCT | ATTTCTCCCT | TTGGAAGAGA | CGCCCTCACC | CATGCTAGGT | GTTCCTCTCC | 18340 |
| | TGGGTGAGGA | GCGTGGCCGT | GTGGCAACCC | CGGGACCTTA | GGCTTATTTA | TTTGTTTAAA | AACATTCCTG | 18410 |
| | GCCTGGCTCT | CGTTGTGTTT | AAATGGGGAA | AAGACATCCC | ACCTCAGCAG | AGTTACTGAG | AGGCTGAAAC | 18480 |
| 35 | CGGGGTGCTG | GCTTGATGTT | TGTGATCTCA | GGTCAATCCA | GAACTGGGCT | AGGAAGCTCAG | TGAGACCAAG | 18550 |
| | TACATGGGGG | GCTCAGGCAG | TGGGTGAGAT | GAGGTACACG | GGGGGCTCAG | GCAGTGGGTG | AGGCCAGGTA | 18620 |
| | CATGGGGGGC | TCAGGCACCT | GGTGAGATGA | GGTACACGGG | GGGCTCAGGC | AGAGGGTCAG | ACCAGGTATA | 18690 |
| | CGGGGGCTCT | GATCACACGC | ACATATGAGC | ACATGTGCAC | ATGTGCTGTT | TCATGTGAGC | CAGGTCTGTG | 18760 |
| | CACACCTGCT | CCAAAGTCCC | AGGAAGCTGA | GAGGCCAAAG | ATGGAGGCTG | ACAGGGCTGG | CGCGGTGGCT | 18830 |
| 40 | CACACCTGCA | GTCCAGCAGC | TTTGGGAGCG | GAGCCGAGA | GGATCCCTTG | AGCCAGGAG | TTTAAGACCA | 18900 |
| | GCCTGAGCAA | CATAGTAGAA | CCCCATCTCT | ATGAAAATAA | AAAAACAAAA | TTAGCTGAAC | ATGGTGGTGT | 18970 |
| | CGCGCTGTAG | TTCCAATACT | TGGGAGGCTG | AAAGTGGGAG | ATACCTTGAG | CCAGGAGGTT | GGAGCTGCA | 19040 |
| | GTGAGCTGAG | ATTGCACCAC | TGTACTGCAG | CCTGGGTGAC | TGATATACGA | CCCATCTCAA | CAACAACAAA | 19110 |
| | GAAAGACTGAC | AATGTCAGTT | TCTTGAAAG | AAACATTTAG | TAGGAACCTA | ACCTACACAC | AGAAGCCAA | 19180 |
| 45 | TCGGTGTCTC | GGTGTCACTG | AGATGAGATG | ATGGGTCCTC | ACACCATCAC | CCAGACACCA | GGGTTTATGC | 19250 |
| | ACCACAGTGG | CGGCTGGCTG | AAGAGGAGTG | CGCAGGACGT | TGATATACGA | TGACATCAAG | TGTGTCTGAC | 19320 |
| | GAAAGGCGAG | ATTCATGATA | AGTACCTGCT | GGTACACAAG | GAACAATGGA | TAAACTGGAA | ACCTTAGAGG | 19390 |
| | CCTTCCCCGA | ACAGGGGGTA | ATCAGAAGCC | AGCATGGGGG | GCTGGCATCT | AGGATGGAGC | TGCTTACGCC | 19460 |
| 50 | TCACATGCGG | TGTTCATACA | GATGGTGACC | AGAAACCGAC | TGTACTGTG | CACACACAGA | CACGCAGCTA | 19530 |
| | CTCGCACACA | CAAGCACACA | CACAGACATG | CATGCATGCA | TCCGTGTGTG | TGCACCTGTG | CCCATGAGGA | 19600 |
| | AATCCATGCA | TGTGCATTTA | TGCAGCAGCA | CAGGCACCGG | TGGGCCCCAT | CCCAACCTCA | CGAGCACCGT | 19670 |
| | CTGATTAGGA | GGCCTTTTCT | CTGACGTGT | CCGCATCTGT | CTCAGGTTTC | ACGCATGTGT | GCTGCAGCTC | 19740 |
| | CCATTTTCATC | AGCAAGTTTG | GAAGAACCCC | ACATTTTTC | TGCGCGTCAT | CTCTGACACG | GCCTCCCTCT | 19810 |
| | GCTACTCCAT | CCTGAAAGCC | AAGAAGCGAG | GTATGTGCAG | GTGCTGGGCC | TCAGTGGGAG | CAGTGCCTGC | 19880 |
| 55 | CTGCTGGTGT | TAGTGTGTCA | GGAGACTGAG | TGAACTTGGG | TCTTAGGAAG | TCTTACCCCT | TTTCGATCTA | 19950 |
| | GGAAAGTGGT | TAACCCAAACC | ACTGTTCAGGC | TCGTCTGCCC | GCCCTCTCGT | GGGGTGAGCA | GAGCACCTGA | 20020 |
| | TGGAAGGGAG | AGGAGCTGTC | TGGGAGCTGC | CATCTTTCCT | ACCTTGTCTT | GCTTGGGGAA | CGCTGGGGGG | 20090 |
| | GCCTGGTCTC | TCTGTTTTGC | CCCATGGTGG | GATTTGGGGG | GCCTGGCCTC | TCCTGTTTGC | CTGTGGTGG | 20160 |
| | GATTGGGCTG | TCTCCCGTCC | ATGGCACTTA | GGGCCCTTGT | GCAAAACCCAG | GCCAAGGGCT | TAGGAGGAGG | 20230 |
| 60 | CCAGGGCCAG | GCTACCCAC | CCCTCTCAGG | AGCAGAGGCC | CGGTATACAC | ACGACAGAGC | CCCCGCCCGT | 20300 |
| | CCTCTGGTTC | CCAGTACCGG | TCTCTTGCCC | CTGGACAGCT | TGTCCAGCAT | CAGGAGGTTT | TTCTGATCCG | 20370 |
| | CTGAAATTCA | AGCCATGTG | AACCTGCGGT | CCTGAGCTTA | ACAGCTTCTA | CTTTCTGTTC | TTTCTGTGTT | 20440 |
| | GTGGAAATTT | CACCTGGAGA | AGCGGAAGAA | AACATTTCTG | TCGTGACTCT | TGCGGTGTCT | GGGTGGGGAG | 20510 |
| | AGCCAGAGAT | GGAGCCACCC | CGCAGACGCT | CGGGTGTGGG | CAGCTTTCCG | GTGTTCTCTG | GGAGGGGAGC | 20580 |
| 65 | TGGGCTGGGC | CTGTGACTCC | TCAGCCTCTG | TTTTCCCCCA | GGGATGTGCG | TGGGGGGCAA | GGGCGCCGCC | 20650 |
| | GGCCCTCTGC | CTCCGAGGCG | CGTGCAGTGG | CTGTGCCACC | AAGCATTTCT | CTCTAAGCTG | ACTCGACACT | 20720 |
| | GTGTACACTA | CGTGCCACTC | CTGGGCTCAC | GTGGCAGAGG | CAAGTGTGGG | TGGAGGCCAG | TGCGGGCCCC | 20790 |
| | ACCTGCCACG | GGGTCACTCT | TGAACGCCCT | GTGTGGGGCG | AGCAGCCTCA | GATGCTGCTG | AAGTGCAGAC | 20860 |
| | GCCCCGGGCG | CTGACCTTGG | GGGCTCTGGG | CCACGCTGGC | AGCCCTATGT | GATTAACAGC | TGGTGTCCCC | 20930 |
| 70 | AGGCCACGGA | GCTTGGGAGG | TGCCCCAACT | TCTTGAACCC | CTGCTTCCCA | TCTCAGGGCG | GATGGCTCCC | 21000 |
| | CAGCCTTGGG | AGCCTTCTGA | CCCTTGACCT | GTGTCTCTCT | ACAGCCTCTT | CCCTGGCTGC | TGCCCTGAGC | 21070 |
| | TCTTGGGGTG | CTGAGCAAGT | TCTCTCCCCG | CCCCCGCCGT | CCAGCGCTAC | TGGGCTGCCT | GTCTGCTCGC | 21140 |
| | CCGGTGGGAG | GGGTGTCTGT | CCCTTCACTG | AGGTTTCCAC | CAGCCAGGCG | CACAGGTTGC | AGGCCCTGCC | 21210 |
| | TGCCCGGCCA | CCACACGCTC | CTAGGAGGGT | TGGAGGATGC | CACCTCTGGC | CTCTTCTGGA | ACGGAGTCTG | 21280 |
| | ATTTTGGCCC | CGCAGCCGAG | ACCGCAGTGA | GTCCGAAGCT | CCCCGGGAGC | AGCCGTCAGT | CCCTGGAGGC | 21350 |
| 75 | CGGAGGCCAC | CGGCACCTGC | CTTCAGACTG | CAAGACCATC | CTGGAAGTAT | AGCCACCCCG | CCACAGCCAG | 21420 |
| | GCCGAGAGCA | GACACCAGCA | GCCCTGTCTC | GCCGGGCTCT | ACGTTCCAGG | GAGGGAGGGG | CGGCCACAC | 21490 |
| | CCAGGGCCCG | ACCGCTGGGA | GTCGTGAGCC | TGAGTGAAGT | TTTTGGCCGAG | GCCGTGCATG | CCGGCTGAAG | 21560 |
| | GCTGAGTGTG | CGGCTGAGGC | CTGAGCGAGT | GTCGCGCAA | GGGCTGAGTG | TCCAGCACAC | CTGCCGTCTT | 21630 |

| | | | | | | | | |
|----|-------------|------------|------------|-------------|------------|-------------|------------|-------|
| | CACTTCCCCA | CAGGCTGGCG | CTCGGCTCCA | CCCCAGGGCC | AGCTTTTCCT | CACCAGGAGC | CCGGCTTCCA | 21700 |
| | CTCCCCACAT | AGGAATAGTC | CATCCCCAGA | TTCCGCATTG | TTCACCCCTC | GCCCTGCCCT | CCTTTGCCTT | 21770 |
| | CCACCCCCAC | CATCCAGGTG | GAGACCCTGA | GAAGGACCCT | GGGAGCTCTG | GGAATTGGA | GTGACCAAAG | 21840 |
| | GTGTGCCCTG | TACACAGGCG | AGGACCCTGC | ACCTGGATGG | GGGTCCCTGT | GGGTCAAATT | GGGGGGAGGT | 21910 |
| 5 | GCTGTGGGAG | TAAAATACTG | AATATATGAG | TTTTTCAGTT | TTGAAAAAAA | TCTCATGTTT | GAATCCTAAT | 21980 |
| | GTGCACTGCA | TAGACACCAC | TGTATGCAAT | TACAGAAGCC | TGTGAGTGAA | CGGGGTGGTG | GTCACTGCGG | 22050 |
| | GCCCCATGGC | TGGCTGTGCA | TTTACGGAAG | TCTATGAGTG | AATGGGGTTG | TGGTCAGTGC | GGGCCCATGG | 22120 |
| | CCTGGCTGGG | CCTGGGAGGT | TTCTGATGCT | GTGAGGCAGG | AGGGGAAGGA | GGGTAGGGGA | TAGACAGTGG | 22190 |
| | GAGCCCCAC | CCTGGAAGAC | ATAACAGTAA | GTCCAGGCC | GAAGGCAGC | AGGGATGCTG | GGGGCCCAGC | 22260 |
| 10 | TTGGGCGGCG | GGGATGATGG | AGGGCCTGGC | CAGGGTGGCA | GGGATGATGG | GGGCCCCAGC | TGGGGTGGCA | 22330 |
| | GGGGTGATGG | GGGGGGCTGG | TCTGGGTGGC | GGGGAAGATG | GGGAAGCCTG | GCTGGGCCCC | CTCCTCCCCT | 22400 |
| | GCCTCCACCC | TGCAGCCGTG | GATCCGGATG | TGCTTCCCTC | GTGCACATCC | TCTGGGCCAT | CAGCTTTCAT | 22470 |
| | GGAGTGGGG | GGCAGGGGCA | TGACACCATC | CTGTATAAAA | TCCAGGATTC | CTCCTCCTGA | ACGCCCAAC | 22540 |
| | TCAGGTTGAA | AGTCACATTC | CGCCTCTGGC | CATTCTCTTA | AGAGTAGACC | AGGATTCTGA | TCTCTGAAGG | 22610 |
| 15 | GTGGGTAGGG | TGGGGCAGTG | GAGGGTGTGG | ACACAGGAGG | CTTCAGGGTG | GGGCTGGTGA | TGCTCTCTCA | 22680 |
| | TCCTCTTATC | ATCTCCAGT | CTCATCTCTC | ATCCTCTTAT | CATCTCCCAG | TCTCATCTGT | CTTCCTCTTA | 22750 |
| | TCTCCAGTCT | TCATCTGTCA | TCCTCTTACC | ATCTCCAGT | CTCATCTCTT | ATCCTCTTAT | CTCCTAGTCT | 22820 |
| | CATCCAGACT | TACCTCCCAG | GGCGGGTGCC | AGGCTCGCAC | TGGAGCTGGA | CATACGTCTC | TCCTCAGGCA | 22890 |
| | GAAGGAAGTG | GAAGGATTGC | AGAGAACAGG | AGGGGCGGCT | CAGAGGGACG | CAGTCTTGGG | GTGAAGAAAC | 22960 |
| 20 | AGCCCTCTCT | CAGAAGTTGG | CTTGGGCCAC | ACGAAACCGA | GGGCCCTGCG | TGAGTGGCTC | CAGAGCCTTC | 23030 |
| | CAGCAGGTCC | CTGGTGGGGC | CTTATGGTAT | GGCCGGGTCC | TACTGAGTGC | ACCTTGGACA | GGGCTTCTGG | 23100 |
| | TTTGAGTGCA | GCCCGGACGT | GCCTGGTGTC | GGGGTGGGGG | CTTATGGCCA | CTGGATATGG | CGTCATTTAT | 23170 |
| | TGCTGCTGCT | TCAGAGAATG | TCTGAGTGAC | CGAGCCTAAT | GTGTATGGTG | GGCCCAAGTG | CACAGACTGT | 23240 |
| 25 | GTCTGTAATG | CACCTCTGGT | CCTGGAGCCC | CCGTATAGGA | GCTGTGAGGA | AGGAGGGGCT | CTTGGCAGCC | 23310 |
| | GGCCTGGGGG | CGCCTTTGCC | CTGCAAACTG | GAAGGGAGCG | GCCCCGGGCG | CCGTGGGCGG | ACGACCTCAA | 23380 |
| | GTGAGAGGTT | GGACAGAACA | GGGCGGGGAC | TTCCCAGGAG | CAGAGGCCGC | TGCTCAGGCA | CACCTGGGTT | 23450 |
| | TGAATCACAG | ACCAACaGGT | CAGGCCATTG | TTCAGCTATC | CATCTTCTAC | AAAGCTCCAG | ATTCTGTGTT | 23520 |
| | CTCCGGGTGT | TTTTTGTGTA | AATTTTACTC | AGGATTACTT | ATATTTTTTG | CTAAAGTATT | AGACCCTTAA | 23590 |
| 30 | AAAAGGTATT | TGCTTTGATA | TGGCTTAACT | CACCTAAGCAC | CTACTTTATT | TGCTGTGTTT | TATTTATTAT | 23660 |
| | TATTTATTAT | ATTAGAGATG | GTGTCTACTC | TGTCACCCAG | GTTGTAGTAG | CAGTGGCACA | GTCTGGGCTC | 23730 |
| | GCTGTAGCCG | CAAACCCCCA | GGCTCAAGTG | ATCCTCCGGC | CTCAGCTTCC | CAGAGTGCTG | GGATTACAGG | 23800 |
| | TGTGAGCCAC | TGCCCTTGCC | TGGCACTTTT | AAAAACCACT | ATGTAAGGTC | AGGTCCAGTG | GCTTCCACAC | 23870 |
| | CTGTCTATCC | AGTAGTTTGG | GAAGCCGAGG | CAGAAGGATT | GTCTGAGGCC | AGGAGTTTGA | GACCAGCATG | 23940 |
| 35 | GGTAACATAG | GGAGACCCCA | TCTCTACAAA | AAATGCAAAA | AGTTATCCGG | GCGTGGGGTC | CAGCATCTGT | 24010 |
| | AGTCCCAGCT | GCTCGGGAGG | CTGAGTGCGA | GGATCGCTTG | AGCCCCGGAG | GTCTATGGCTG | CAGTGAGCTG | 24080 |
| | TGATTGTACC | ATCGCACTCC | AGCCTGGGCA | ACAGAGTGAG | ACCTGTCTCT | AAAAAATAAA | AAAAAATAAG | 24150 |
| | AAGGAGAAGG | AGAAGAGAAG | AAGAAGGAAG | AAGGAAAGAG | AAGAAGAAGG | AAGAAGGAAG | AAAGAAGGAG | 24220 |
| | AAGGAGGCC | GCTAGGTGCT | AGGTAGACTG | TCAAATCTCA | GAGCAAAATG | AAAAATAACAA | AGTTTAAAGG | 24290 |
| 40 | GGAAAGAAAA | ACCCAGCTC | TTTGGACTTC | CTTAGGCCTG | AACTTCATCT | CAAGCAGCTT | CCTTCCACAG | 24360 |
| | ACAAGCGTGT | ATGGAGCGAG | TGAGTTCAAA | GCAGAAAGGG | AGGAGAAGCA | GGCAAGGGTG | GAGGCTGTGG | 24430 |
| | GTGACACCAG | CCAGGACCCC | TGAAAGGGAG | TGGTTGTTTT | CCTGCCTCAG | CCCCACGCTC | CTGCCGGTCC | 24500 |
| | TGCACCTGCT | GTAACCGTCG | ATGTTGGTGC | CAGGTGGCCA | CCTGGGAAGG | ATGCTGTGCA | GGGGGCTTGC | 24570 |
| | CAAACCTTGG | TGGGTTTCAG | AAGCCCCAGG | CACCTGTGGC | AGGCACAATT | ACAGCCCCTC | CCCAAGATG | 24640 |
| 45 | CCCACGTCTCT | TCTCCTGGAA | CCTGTGAATG | TGTCACCCGC | AAGGCAGAGG | CTGGTGAAGG | CTGCAGGTGG | 24710 |
| | AATCAGGGCT | GCCAGTCAGC | CGATCTTAAG | GTCACTCTGG | ATTATCTGGT | GGGCTTGATA | TGGCCACAAG | 24780 |
| | GGTCCCTAGA | AGTGAGAGAG | GGAGGCAGGG | GAGAGTCAGA | GAGGGGACGT | GAGAAGGACC | ACTGGCCACT | 24850 |
| | GCTGGCTTTG | AGATGGAGGA | GGGGGTCCCC | AGCCAAGGAA | TGGGGGCAGC | CGCTCCATGC | TGGAAAAGCA | 24920 |
| | AGCAATCCTC | CCCGGTCCTG | AGGGCACACG | GCCTCGATT | TCAGGCCAGT | GGGACCTGTT | GGGACCTGTT | 24990 |
| | TCAGCTTTCC | GGCCTCCAGA | GCTGTAAGAT | GATGCGTTTG | TGTTAGCCCA | CTAAGCTGCA | GTGATTCTGC | 25060 |
| 50 | ACAGCAGCAA | ATGGAATAGC | AGTACAGGGA | AATGAATACA | GGGACAGTTC | TCAGAGTGAC | TCTCAGCCCA | 25130 |
| | CCCCCTGGG | | | | | | | 25138 |

Example 5

- 55 Comparison of the above-described genomic hTC sequence and the sequence of the hTC cDNA (Fig. 6: corresponding to SEQ ID NO 2) made it possible to elucidate the exon-intron structure of the hTC gene. The genomic organization of the hTC gene is illustrated diagrammatically in Fig. 7. The coding region of the hTC gene is composed of 16 exons which vary in size between 62 bp and 1354 bp (see Table 1).
- 60 Exon 1 contains the translation start codon ATG. The translation stop codon TGA and the 3'-untranslated region lie on exon 16 (Fig. 8). No possible polyadenylation signal (AATAAA) was found either in exon 16 or in the 3195 bp of the following

3'-flanking region. The exon-intron transitions were determined on the basis of the consensus sequence

| | | 5'-Exon | | | Intron | | | | | 3'-Exon | | |
|---|---------------|---------|----|----|--------|-----|-----|-----|----|---------|-----|------------|
| 5 | Pre-mRNA | A/C | A | G | | G | T | A/G | A | ... | N C | A G G |
| | Frequency (%) | 70 | 60 | 80 | | 100 | 100 | 95 | 70 | | 80 | 100 100 60 |

and listed in Table 1. With the exception of the 5' splice site between exon 15 and
intron 15, all the exon-intron transitions are in accord with the published (Shapiro
and Senapathy, 1987) splice consensus sequence. The sizes of the introns are
between 104 bp and 8616 bp. Since only part of intron 6 was isolated, it is not
possible to determine the precise length of the hTC gene. Based on the part sequence
of ~4660 bp, which was obtained from intron 6, the minimum size of the hTERT
gene is 37 kb.

Introns 1-5 and the 5' region of intron 6, are contained in contig 1:

Intron 1: bp 11493-11596 (SEQ ID NO 4);

Intron 2: bp 12951-21566 (SEQ ID NO 5);

Intron 3: bp 21763-23851 (SEQ ID NO 6);

5 Intron 4: bp 24033-24719 (SEQ ID NO 7);

Intron 5: bp 24900-25393 (SEQ ID NO 8);

5' region of intron 6: bp 25550-26414 (SEQ ID NO 9).

10 The 3' region of intron 6, and introns 7-15, are located in contig 2 at the following positions:

3' region of intron 6: bp 1-3782 (SEQ ID NO 10);

Intron 7: bp 3879-4858 (SEQ ID NO 11);

Intron 8: bp 4945-7429 (SEQ ID NO 12);

Intron 9: bp 7544-9527 (SEQ ID NO 13);

15 Intron 10: bp 9600-11470 (SEQ ID NO 14);

Intron 11: bp 11660-15460 (SEQ ID NO 15);

Intron 12: bp 15588-16467 (SEQ ID NO 16);

Intron 13: bp 16530-19715 (SEQ ID NO 17);

Intron 14: 19841-20621 (SEQ ID NO 18);

20 Intron 15: 20760-21295 (SEQ ID NO 19).

The 3'-untranscribed region is also located in contig 2 at position 21960-25138 (SEQ ID NO 20).

25 The individual sequences of the abovementioned introns are as follows:

00T250" 91223560

Intron 1 (SEQ ID NO 4)

GTGGGCCTCCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGG
CGACTCAGGGCGCTTCCCCCGCAG

5 Intron 2 (SEQ ID NO 5)

GTGAGGAGGTGGTGGCCGTCGAGGGCCCAGGCCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCC
CTGGTCCTCCTGTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTGAGTGACACGGTGATCTCTGCC
TCTGCTCTCCCTCCTGTCCAGTTTGCATAAACTTACGAGGTTACCTTCACGTTTTGATGGACACGCGGTTTCCAGGCGC
CGAGGCCAGAGCAGTGAACAGAGGAGGCTGGGCGCGGCAGTGAGCCGGGTTGCCGGAATGGGGAGAAGTGTCTGGAAG
10 CACAGACGCTCTGGCGAGGGTGCTGCAGGTTACCTATAATCCTCTTCGCAATTTCAAGGTGGGAATGAGAGGTGGGGA
CGAGAACCCCTCTTCTGGGGTGGGAGGTAAGGGTTTTGCAGGTGCACGTGGTCAGCCAATATGCAGGTTTGTGTTTA
AGATTTAATTGTGTGTTGACGGCCAGGTGCCTGGCTCACGCCGGAATCCCAGCACTTTGGGAAGCTGAGGCAGGTGGA
TCACCTGAGGTGAGGAGTTTGAGACCAGCCTGACCAACATGGTGAAACCCTATCTGTACTAAAAATACAAAATTAGCTG
GGCATGGTGGTGTGTGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCGGAGGC
15 TGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCGACAAGAGTGAACTCTGTCTTTAAAAAAGTGT
CGTTGATTGTGCCAGGACAGGGTAGAGGGAGGGAGATAAGACTGTTCTCCAGCACAGATCCTGGTCCCATCTTTAGGTAT
GAAGAGGGCCACATGGGAGCAGAGGACAGCAGATGGCTCCACCTGCTGAGGAAGGGACAGTGTTTGTGGGTGTTAGGGG
ATGGTGCTGCTGGGCCCTGCCGTGTCCCAACCTGTTTTCTGGATTTGATGTTGAGGAACCTCCGCTCCAGCCCCCTTT
TGGCTCCCAGTGCTCCAGGCCCTACCGTGCCAGCTAGAAGAAGTCCCGATTTACCCCCCTCCCCACAACTCCCAAGAC
20 ATGTAAGACTTCCGGCCATGCAGACAAGGAGGGTGACCTTCTTGGGGCTCTTTTTTTCTTTTTTTCTTTTTATGGTGGC
AAAAGTCATATAACATGAGATTGGCACTCCTAACACCGTTTTCTGTGTACAGTGACAGAATTGCTAACTCGGCGGTGTTTA
CAGCAGGTTGCTTGAAATGCTGCGTCTTGCGTGACTGGAAGTCCCTACCCATCGAACGGCAGCTGCCTCACACCTGCTGC
GGCTCAGGTGGACCACGCCGAGTCAGATAAGCGTCATGCAACCCAGTTTTGCTTTTTGTGCTCCAGCTTCCTTCGTTGAG
GAGAGTTTGAGTTCTCTGATCAGGACTCTGCCTGTCTGCTGTTCTCTGACTTCAGATGAGGTACAATCTGCCCTGG
25 CTTATGCAGGGAGTGAGGCGTGGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGGCGTTGCCCCCAGGTGTCCCT
GTCACGTGTAGGGTGAGTGAGGCGCGGCCCCGGGTGTCCCTGTCCCGTGACGCTGATTGAGGTGTGGCCCCGGGTGT
CCCTGTACGTGTAGGGTGAGTGAGGCGCCATCCCCGGGTGTCCCTGTACGTGTAGGGTGAGTGAGGCGTGGTCCCCGG
GTGTCCCTGTCCCGTGACAGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGGCGCGGTCC
CCGGGTGTCCCTCTCAGGTGTAGGGTGAGTGAGGCGCGGCCCCAGGGTGTCCTGTACGTGTAGGGTGAGTGAGGCACC
30 GTCCCTGGGTGTCCCTCCCAGGTATAGGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGG
CGCGGCCCCGGGTGTCCCTCTCAGGTGCAGGGTGAGTGAGGCGCTGTCCCTGGGTGTCCCTGTCTCGTGTAGGGTGAGT
GAGGCTCTGTCCCCAGGTGTCTTGGCGTTTGCTCACTTGAGCTTGCTCCTGAATGTTTGCTCTTTCTATAGCCACAGCT
GCGCCGGTTGCCATTGCCTGGGTAGATGGTGCAGGCGCAGTGCTGGTCCCCAAGCCTATCTTTTCTGATGCTCGGCTCT
TCTTGGTCACCTCTCCGTTCCATTTTGCTACGGGGACACGGGACTGCAGGCTCTCGCCTCCCGCGTGCCAGGCACTGCAG
35 CCACAGCTTCAGGTCCGCTTGCCCTCTGTTGGGCCTGGCTTGCTCACCACGTGCCCGCCACATGCATGCTGCCAATACTCC
TCTCCCAGCTTGCTCATGCCGAGGCTGGACTCTGGGCTGCCTGTGTCTGTGCCACGTGTGTGAGGACATCCCAGAA
AGGGTTCTCTGTGCCCTGAAGGAAAGCAAGTCACCCAGCCCCCTCACTTGCTCTGTTTTCTCCCAAGCTGCCCTCTGC
TTGGCCCCCTTGGGTGGGTGGCAACGCTTGTCACCTTATTTCTGGGCACCTGCCGCTCATTGCTTAGGCTGGGCTCTGCCT
CCAGTCGCCCCCTCACATGGATTGACGTCCAGCCACAGGTGGAGTGCTCTGTCTGTCTCCTGCTCTGAGACCCACGTG
40 GAGGGCGCGGTGTCTCGCCAGCCTTCGTGCACTTCCCTCTTGGGTCTTAGTTTTGAATTTCACTGATTTACCTCTGACG
TTTCTATCTCTCCATTGTATGCTTTTTCTTGGTTTTATTCTTTTCATTCCTTTCTAGCTTCTTAGTTAGTCATGCTTTT
CCTCTAAGTGCTGCCTTACCTGCACCTGTGTTTTGATGTGAAGTAATCTCAACATCAGCCACTTTCAAGTGTCTTAA
ATACTTCAAAGTGTTAATACTTCTTTTAAGTATTCTTATTTCTGTGATTTTTCTTTGTGCACGCTGTGTTTTGACGTGA
AATCATTTTGATATCAGTGACTTTTAAGTATTCTTTAGCTTATTCTGTGATTTCTTTGAGCAGTGAGTTATTTGAACACT
45 GTTTATGTTCAAGATATGTAGAGTATCAAGATACGTAGAGTATTTTAAGTTATCATTTTATTATTGATTTCTAACTCAGT
TGTGTAGTGGTCTGTAATAACCAATTATTTGAAGTTTGGGAGCCTTGCTTTGTGATCTAGTGTGTGCATGGTTTCCAG
AACTGTCCATTGTAATTTGACATCCTGTCAATAGTGGGCATGCATGTTCACTATATCCAGCTTATTAAGGTCCAGTGCA

AAGCTTCTGCTCTCCTCTAGATGCATGAAATTCGAAGGAGGCCATAGTCCCTCACCTGGGGGATGGGTCTGTTCAATT
 TCTTCTCGTTTTGGTAGCATTTATGTGAGGCATTGTTAGGTGCATGCACGTGGTAGAATTTTTATCTTCTGATGAGTGAA
 TCTTTTGGAGACTTCTATGTCTCTAGTAATCTAGTAATTCCTTTTTTAAATTGCTCTTAGTACTGCCACACTGGGCTTCT
 TTTGATTAGTATTTTCTGCTGTGTCTGTTTTCTGCCTTTAATTTATATATATATATATATTTTTTTTTTTTTTTGAGACA
 5 GAGTCTTGGTCTGTGCGCCAGGGTGAGTGCAGTGGTGTGATCACAGGTCAGTGTAACTTTTACCTTCTGGCCTGAGCCGT
 CCTCTCACCTCAGCCTCCTGAGTAGCTGGAAGTGCAGACACGCACCGCTACACCTGGCTAATTTTTAAATTTTTTCTGGA
 GACAGGGTCTTGCTGTGTTGCCAGGCTGGTCTCAAACCTCTGGACTCAAGGGATCCATCTACCTCGGCTTCCCAAAGTG
 CTGAATTACAGGCATGAGCCACCATGTCTGGCCTAATTTTCAACACTTTTATATTCTTATAGTGTGGGTATGTCCTGTTA
 ACAGCATGTAGGTGAATTTCCAATCCAGTCTGACAGTCGTTGTTTAACTGGATAACCTGATTTATTTTCATTTTTTTGTG
 10 ACTAGAGACCCGCTGGTGCACCTCTGATTCTCCACTTGCTGTGTCATGTCTCGTTCCTTGTCTTCTACCACCTCTTG
 GGTTCGCATGTGCGTTTTCTGCGGAGTGTGTGTTGATCCTCTCGTTGCCTCCTGGTCACTGGGCATTTGCTTTTATTTCT
 CTTTGCTTAGTGTTACCCCTGATCTTTTTATTGTGCTGTTGTTGCTTTTGTATTAGAGACAGTCTCACTCTGTCAACCA
 GGCTGGAGTGTAATGGCACAATCTCGGCTCACTGCAACCTCTGCCTCCTCGGTTCAAGCAGTTCTCATTCCTCAACCTCA
 TGAGTAGCTGGGATTACAGGCGCCACCACCACGCCTGGCTAATTTTGTATTTTTTAGTAGAGATAGGCTTTACCATTGT
 15 TGGCCAGGCTGGTCTCAAACCTCCTGACCTCAAGTGATCTGCCCGCCTTGGCCTCCACAGTGCTGGGATTACAGGTGCAA
 GCCACCGTGCCCGGCATACCTTGATCTTTTAAATGAAGTCTGAAACATTGCTACCTTGTCTGAGCAATAAGACCCTT
 AGTGTATTTTAGCTCTGGCCACCCCCAGCCTGTGTGCTGTTTTCCCTGCTGACTTAGTTCTATCTCAGGCATCTTGACA
 CCCCCACAAGCTAAGCATTATTAATATTGTTTTCCGTGTTGAGTGTTTCTGTAGCTTTGCCCCCGCCCTGCTTTTCTCTC
 TTTGTTCCCGTCTGTCTTCTGTCTCAGGCCGCGCGTCTGGGGTCCCTTCTTGTCTTTGCGTGGTTCTTCTGTCTTG
 20 TTATTGCTGGTAAACCCAGCTTTACCTGTGCTGGCCTCCATGGCATCTAGCGACGTCCGGGGACCTCTGCTTATGATGC
 ACAGATGAAGATGTGGAGACTCACGAGGAGGGCGGTCTCTGGCCGTGAGTGTCTGGAGCACCAGTGCCAGCGTTC
 CTTAGCCAGTGAGTGACAGCAACGTCCGCFCGGCCTGGGTTCAGCCTGGAAAACCCAGGCATGTGCGGGTCTGGTGGCT
 CCGCGGTGTGAGTTTGAATCGCGCAAACCTGCGGTGTGGCGCCAGCTCTGACGGTGCTGCCTGGCGGGGGAGTGTCTG
 CTTCTCCCTTCTGCTTGGGAACAGGACAAAGGATGAGGCTCCGAGCCGTTGTGCGCCAACACGAGCATGACGTGAGCC
 25 ATGTGGATAATTTTAAATTTCTAGGCTGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGG
 TGGATCACGAGGTCAGGAGTCTGAGACCATCCTGGCCAACATGATGAAACCCCATCTGTACTAAAAACACAAAAATTAGC
 TGGGCGTGGTGGCGGGTGCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGTTGGAA
 GTTGACAGTGAGCCGACATTGCACCACTGCACTCCAGCCTGGCAACACAGCGAGACTCTGTCTCAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAATTCTAGTAGCCACATTAATAAAGTAAAAAAGAAAAGGTGAAATTAATGTAATAATAGATTTTACTGAA
 30 GCCCAGCATGTCCACACCTCATCATTTTAGGGTGTTATTGGTGGGAGCATCACTCACAGGACATTTGACATTTTTTGTAGC
 TTTGTCTGCGGGATCCCCTGTGTAGGTCCCCTGCGTGGCCATCTCGGCCTGGACCTGCTGGGCTTCCCATGGCCATGGCT
 GTTGTACCAGATGGTGCAGGTCCGGGATGAGGTGCGCCAGGCCCTCAGTGAGCTGGATGTGCAGTGTCCGGATGGTGCACG
 TCTGGGATGAGGTGCGCAGGCCCTGCTGTGAGCTGGATGTGTGGTGTCTGGATGGTGCAGGTGAGGGGTGAGGTCTCCAG
 GCCCTCGGTGAGCTGGAGGTATGGAGTCCGGATGATGCAGGTCCGGGGTGGAGTTCGCCAGGCCCTGCTGTGAGCTGGATG
 35 TGTGGTGTCTGGATGGTGCAGGTGAGGGGTGAGGTCTCCAGGCCCTCGGTAAGCTGGAGGTATGGAGTCCGGATGATGCA
 GGTCCGGGGTGGAGTTCGCCAGGCCCTGCTGTGAGCTGGATGTGTGGTGTCTGGATGGTGCAGGTCTGGGGTGGAGTCAAC
 AGGCCCTGCGGTGAGCTGGGTGTGCGGTGTCTGGATGGTGCAGGTCTGGAGTGGAGTTCGCCAGACGGTGCCAGACCATGC
 GGTGAGCTGGATATGCGGTGTCCGGATGGTGCAGGTCTGGGGTGGAGTTGCCAGGCCCTGCTGTGAGTTGGATGTGGGGT
 GTCCGGATGCTGCAGGTCCGGTGTGAGGTACCAGGCCCTGCTGTGAGCTGGATGTGTGGTGTCTGGATGGTGCAGGTCT
 40 GGGGTGAAGGTGCGCAGGCCCTGCTTGTGAGCTGGATGTGTGGTGTCTGGATGGTGCAGGTCTGGAGTGGAGTTCGCCAG
 GCCCTCGGTGAGCTGGATGTGCAGTGTCCAGATGGTGCAGGTCCGGGGTGGAGTTCGCCAGACCCCTGCGGTGAGCTGGATG
 TGCGGTGTCTGGATGGTGCAGGTCTGGAGTGGAGTTCGCCAGGCCCTCGGTGAGCTGGATGTATGGAGTCCGGATGGTGCC
 GGTCCGGGGTGGAGTTCGCCAGGCCCTGCTGTGAGCTGGATGTGCGGTGTCTGGATGGTACAGGTCTGGAGTGGAGTTCGCC
 AGACCCTGCTGTGAGCTGGATATGCGGTGTCCGGATGGTGCAGGTGAGGGGTGGAGTCTCCAGGCCCTCGGTGAGCTGGA
 45 GGTATGGAGTCCGGATGATGCAGGTCCGGGGTGGAGTTCGCCAGGCCCTGCTGTGAACCTGGATGTGCGGCGTCTGGATGGT
 GCAGGTCTGGGGTGTGGTTCGCCAGGCCCTCGGTGAGCTGGAGGTATGGAGTCCGGATGATGCAGGTCCGGGGTGGAGTTCG
 CCAGGCCCTGCTGTGAGCTGGATGTGCGGCGTCTGGATGGTGCAGGTCTGGGGTGTGGTTCGCCAGGCCCTCGGTGAGCT

5 GAGGTATGGAGTCCGGATGATGCAGGTCCGGGGTGAGGTTGCCAGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATG
GTGCAGTCCGGGGTGAGGTCGCCAGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCTGGGGTGAGGT
CACCAGGCCCTGCGGTGAGCTGGTTGTGCGGTGTCCGGTTGCTGCAGGTCCGGGGTGAGTTCGCCAGGCCCTCGGTGAGC
TGGATGTGCGGTGTCCCCGTGTCCGGATGGTGCAGGTCCAGGGTGAGGTGCGTAGGCCCTTGGTGGGCTGGATGTGCCGT
10 GTCCGGATGGTGCAGGTCTGGGGTGAGGTGCCAGGCCCTTGGTGAGCTGGATGTGCGGTGTCTGCATGGTGCAGGTCTG
GGGTGAGGTGCCAGGCCCTTGGTGGGCTGGATGTGTGGTGTCCGGATGGTGCAGGTCCGGCGTGAGGTGCCAGGCCCT
GCTGTGAGCTGGATGTGCGGTGTCTGGATGGTGCAGGTCCGGGGTGAGGTAGCCAAGGCCCTTCGGTGAGCTGGATGTGGG
GTGTCCGGATGGTGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCGGTAGCTGGATATGCGGTGTCCGGATGGTGCAGGT
CCGGGGTGAGGTACCAGGCCCTGCGGTAGCTGGATGTGCGGTGTCTGGATGGTGCAGGTCCGGGGTGAGGTGCCAGG
15 CCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCAGTGAGCTGGATG
TGCTGTATCCGGATGGTGCAGGTCTGGCGTGAGGTGCCAGGCCCTGCGGTTAGCTGGATATGCGGTGTCCGGATGGTGCA
GGTCCGGGGTGAGGTACCAGGCCCTGCGGTAGCTGGATGTGCGGTGTCCGGATGGTGCAGGTCTGGGGTGAGGTGCC
AGGCCCTGCTGTGAGCTGGATGTGCTGTATCCGGATGGTGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCGGTGAGCTGG
ATGTGCTGTATCCGGATGGTGCAGGTCTGGCGTGAGGTGCCAGGCCCTGCGGTGAGCTGGATGTGCAGTGTACGGATGG
20 TGCAGGTCCGGGGTGAGGTGCCAGGCCCTGCGGTGGGCTGTATGTGTGTGTCTGGATGGTGCAGGTCCGGGGTGAGTT
CGCCAGGCCCTGCGGTGAGCTGGATGTGTGGTGTCTGGATGCTGCAGGTCCGGGGTGAGTTCGCCAGGCCCTCGGTGAGC
TGGATATGCGGTGTCCCCGTGTCCGAATGGTGCAGGTCCAGGGTGAGGTGCCAGGCCCTTGGTGGGCTGGATGTGCCGT
GTCCGGATGGTGCAGGTCTGGGGTGAGGTGCCAGGCCCTTGGTGAGCTGGATGTGCGGTGTCCGGATGGTGCAGGTCCG
GGGTGAGGTACCAGGCCCTCGGTGATCTGGATGTGGCATGTCTTCTCGTTTAAG

Intron 3 (SEQ ID NO 6)

GTACTGTATCCCCACGCCAGGCCCTGTCTTCTCGAAGTCCTGGAACACCAGCCCCGGCTCAGCATGCGCCTGTCTCCACT
TGCCTGTGCTTCCCTGGCTGTGCAGCTCTGGGCTGGGAGCCAGGGGCCCCGTACAGGCCTGGTCCAAGTGGATTCTGTG
CAAGGCTCTGACTGCCTGGAGCTCACGTTCTCTTACTTGTAAATCAGGAGTTTGTGCCAAGTGGTCTCTAGGGTTTGTA
25 AAGCAGAAGGGATTTAAATTAGATGGAACACTACCACTAGCCTCCTTGCCCTTCCCTGGGATGTGGGTCTGATTCTCTC
TCTCTTTTTTTTTTCTTTTTTGTAGATGGAGTCTCACTCTGTTGCCAGGCTGGAGTGCAGTGGCATAATCTTGGCTCACT
GCAACCTCCACCTCCTGGGTTTAAGCGATTACACAGCCTCAGCCTCCTAAGTAGCTGGGATTACAGGCACCTGCCACCAC
GCCTGGCTAATTTTTGTACTTTTAGGAGAGACGGGGTTTACCATGTTGGCCAGGCTGGTCTCGAACTCATGACCTCAGG
TGATCCACCCACCTTGGCCTCCCAAAGTGTGGGTTTACAGGCTAAGCCACCGTGCCAGCCCCGATTCTCTTTAATT
30 CATGCTGTTCTGTATGAATCTTCAATCTATTGGATTTAGGTCATGAGAGGATAAAATCCACCCACTTGGCGACTCACTG
CAGGGAGCACCTGTGCAGGGAGCACCTGGGGATAGGAGAGTTCCACCATGAGCTAAGTCTAGGTGGCTGCATTGGAATG
GCTGTGAGATTTGTCTGCAATGTTCCGGCTGATGAGAGTGTGAGATTGTGACAGATTCAAGCTGGATTTGCATCAGTGAG
GGACGGGAGCGCTGGTCTGGGAGATGCCAGCCTGGCTGAGCCCAGGCCATGGTATTAGCTTCTCCGTGTCCGCCCAGGC
TGACTGTGGAGGGCTTTAGTCAGAAGATCAGGGCTTCCCCAGCTCCCCTGCACACTCGAGTCCCTGGGGGGCCTTGTGAC
35 ACCCATGCCCCAAATCAGGATGTCTGCAGAGGGAGCTGGCAGCAGACCTCGTCAGAGGTAACACAGCCTCTGGGCTGGG
GACCCCGACGTGGTGTGGGGCCATTTCTTGCATCTGGGGGAGGGTCAGGGCTTTCCTGTGGGAACAAGTTAATACAC
AATGCACCTTACTTAGACTTTACACGTATTTAATGGTGTGCGACCCAACATGGTCATTTGACCAGTATTTTGAAAGAAT
TTAATTGGGGTGACCGGAAGGAGCAGACAGACGTGGTGGTCCCCAAGATGCTCCTTGTCACTACTGGGACTGTTGTTCTG
CCTGGGGGGCCTTGGAGGCCCTCCTCCCTGGACAGGGTACCGTGCTTTTCTACTCTGCTGGGCCTGCGGCCTGCGGTG
40 AGGGCACCAGCTCCGGAGCACCCGCGGCCCCAGTGTCACAGGAGTGCCAGGCTGTACGCCACAGATGCCAGGTCCAGGT
GTGGCCGCTCCAGCCCCCGTCCCCCATGGTGGTTTTGGGGGAAAAGGCCAAGGGCAGAGGTGTGAGGAGACTGGTGGG
CTCATGAGAGCTGATTCTGCTCCTTGGCTGAGCTGCCCTGAGCAGCCTCTCCCGCCCTCTCCATCTGAAGGGATGTGGCT
CTTTCTACCTGGGGGCTCCTGCTGGGGCCAGCCTTGGGCTACCCAGTGGCTGTACCAGAGGGACAGGCATCCTGTGTGG
AGGGGCATGGGTTCACGTGGCCCCAGATGCAGCCTGGGACCAGGCTCCCTGGTGTGATGGTGGGACAGTCACCCTGGGG
45 GTTGACCGCCGACTGGGCTCCCCAGGGTTGACTATAGGACCAGGTGTCCAGGTGCCCTGCAAGTAGAGGGGCTCTCAG
AGGCGTCTGGCTGTCATGGGTGGACGTGGCCCCGGGCATGGCCTTCAGCGTGTGCTGCCGTGGGTGCCCTGAGCCCTCAC
TGAGTCGGTGGGGCTTGTGGCTTCCCGTGAGCTTCCCCCTAGTCTGTTGTCTGGCTGAGCAAGCCTCCTGAGGGGCTCT
CTATTGCAG

5

10

15

20

25

30

35

40

CAGAAGAGTTTCACGTGTGCTGATTTCCCGGCTGTTTCTGCGTAATTGGTGTCTGCTGTTTATCGATGGCCTCCTTCCA
 TTTCCCTTAGGCTTTGTTTATTGTTGTTTTTCCGGCTCCTTGAAGGAAAAGTTTCGATTATGGATGTTTGAACCTTCTTT
 TCTAAACAAGCATCTGAAGTTGCCGTTTTCCCTCTAAAGCAGGGATCCCGAGGCCCTGGCTGTGGAGTGGCACCAGGTCT
 5 GGGGCCCTGTTAGGAACCCGGCGCACAGCGGAGGCTAGGTGGGGTGTGGGGAGCCAGCGTTCCCGCCTGAGCCCCGCCCC
 TCTCAGATCAGCAGTGGCATGCGGTGCTCAGAGGCGCACACACCCTACTGAGAACTGTGCGTGAGAGGGGTCTAGATTCT
 GTGCTCCTTATGGGAATCTAATGCCTGATGATCTGAGGTGGAACCGTTTGCTCCCAAACCATCCCTTCCCCACTGCTG
 TCCTGTGGAAAAATCGTCTTCCACGAAACAGTCCCTGGTACCACAATGGTTGGGGACCCTGTGCTAAAGACCTGCTTCA
 GCAGCCTCTCGTCAGTGTGATATATTGGCTTTTCTGTGTTGAGTCCAGAATAATTACGGATTTCTGTGATGCTTTCCGC
 CGACCTCAGACCCATGGGCTATTTGTGGGCGTGTGCCTGCTCCTGGGTGGGAAGGTGCAGGCCCATGTACCTTCTT
 10 GTTACTGCCCTTCCAGGTTGGTTCTCAGGGTTGAATCGTACTCGATGTGGTTTTAGCCCACGGCCCTGCCGCCAGCTCCTG
 GGGGCTGGGGAACATGCTGAAGCACAGAGTCACCGTGCGCGTCTTTTGATGCCTCACAAGCTCGAGGCCTCCTGTGTCCG
 TGTTAGTGTGTGTACGTGCCTGCTCACATCCTGTCTTGGGACGCAGGGGCTTAGCAGGTCCCGTAGTAAATGACAAGC
 GTCCTGGGGGAGTCTGCAGAATAGGAGGTGGGGGTGCCGGTCTCTCTCCCGCGTCTTCAGACTCTTCTCCTGCCTGTGCT
 GTGGCTGCACCTGCATCCCTGCAATCCCTCCAGCACTGGGCTGGAGAGGCCCGGGAGCTCGAGTGCCACTTGTGCCACGT
 15 GACTGTGGATGGCAGTCCGTACCGGGGTCTGATGTGTGGTGAATGTGGATGGCGGTGGTTCACAGGGGTCTGATGTGTG
 GTGACTGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGATGTGTGGTGAATGTGGATGG
 ATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTG
 GGGTCTGATGTGGTGAATGTGGATGGCAGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATG
 TGGTGAATGTGGATGGCAGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAAT
 20 GTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGG
 CGGTCTGTTGGGTCTGATGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGTGAATCGGTCA
 CAGGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGTGAATCGGTCA
 CAGGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTG
 GGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTG
 25 TCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTG
 GATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGGT
 GACTGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGAT
 GGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCAG
 TCGTGGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTG
 30 TCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGT
 GGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGTGAATCGGTCAAGGGGTCTGATGTGTGGT
 GACTGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGAT
 GGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCAGTCTG
 GTCACAGGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGG
 GGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGATGTGTGGTGAATGTGGATGGCGGTCTGTTGGGTCTGAT
 35 GTGGTGAATGTGGATGGTGAATCGGTCAAGGGGTCTGATGTGTGGTAGCTGCAGGTGGAGTCCAGGTGTGTCTGTAGCT
 ACTTTGCGTCTCGGCCCCCGGCCCCCGTTTCCCAAACAGAAGCTTCCAGGCGCTCTCTGGGCTTCATCCCGCCATCG
 GGCTTGGCCGCAGGTCCACACGTCCTGATCGGAAGAAACAAGTGCCAGCTCTGGCCGGGCAGGCCACATTTGTGGCTC
 ATGCCCTCTCCTCTGCCGCCAG

40 Intron 7 (SEQ ID NO 11)

GTCTGGGCACTGCCCTGCAGGGTTGGGCACGGACTCCAGCAGTGGGTCTCCCTGGGCAATCACTGGGCTCATGACCG
 GACAGACTCTTGGCCCTGSGGGGCAGTGGGGGAATGAGCTGTGATGGGGGCATGATGAGCTGTGTGCCCTTGGCGAAATC
 TGAGCTGGGCCATGCCAGGCTGCGACAGCTGCTGCATTACAGGCACCTGCTCACGTTTACTGCGCGGCCCTCTCTCCAGTT
 CCGCAGTGCCTTTGTTCAATGATTTGCTAAATGTCTTCTCTGCCAGTTTTGATCTTGAGGCCAAAGGAAAGGTGTCCCCCT
 45 CCTTTAGGAGGGCAGGCCATGTTTGAACCGTGTCTGCCAGCTGGCCCCCTCAGTGTGGGTCTGAGGCCAAAGGAAACG
 TGTCCCCCTCTTAGGAGTACGGGGCCGTGTTTGAACACGCCCGCTGAGCGGGCCTCTCAGTGTGGGTCTGTCCACGT

GGCCCTGTGGCCCTTTGCAGATGTGGTCTGTCCACGTGGCCCTGTGGCTCTTTGCAGATGCCTGTTAGCACTTGCTCGGC
TCTAGGGGACAGTCGTGTCCACCGCATGAGGCTCAGAGACCTCTGGGCGAATTTCCCTTGGCTCCCAGGGTGGGGGTGGAG
GTGGCCTGGGCTGCTGGGACCCAGACCCCTGTGCCCGGCAGCTGGGCAGCAACTCCTGGATCACATATGCCATCCGGGCCA
CGGTGGGCTGTGTGGGTGTGAGCCCAGCTGGACCCACAGGTGGCCAGAGGAGACGTTCTGTGTACACACTCTGCCTAA
5 GCCCATGTGTGTCTGCAGAGACTCGGCCCGCCAGCCACGATGGCCCTGCATTCCAGCCCAGCCCCGCACTTCATCACA
AACACTGACCCCAAAGGGACGGAGGGTCTTGCCACGTGGTCTGCCTGTCTCAGCACCCACGGGCTCACTCCCATGTG
TCTCCCGTCTGCTTTCGCAG

Intron 8 (SEQ ID NO 12)

10 GTGAGTCAGGTGGCCAGGTGCCATTGCCCTGCGGGTGGCTGGGCGGGCTGGCAGGGCTTCTGCTCACCTCTCTCCTGCCC
CTTCCCCACTGNCCTTCTGCCCCGGGGCCACCAGAGTCTCCTTTTCTGGCCCCCGCCCCCTCCGGCTCCTGGGCTGCAGGC
TCCCGAGGCCCCGAAACATGGCTCGGCTTGCGGCAGCCGAGCGGAGCAGGTGCCACACGAGGCCTGGAAATGGCAAGC
GGGGTGTGGAGTTGCTCCTGCGTGGAGGACGAGGGGCGGGGGTGTGTCTGGGTGAGGTGTGCGCCGAGCGTTTGAGCCT
GCAGCTTGTGAGTCCAAGTTACTACTGACGCTGGACACCCGGCTCTCACACGCTTGTATCTCTCTCTCCCCATACAAAA
15 GGATTTTATCCGATTCTCATTCCTGTCCCTGTGCTGTGACCCCCGCGAGGGCGGGGCTCTTCTCTCTGTGACTAGATTT
CCCCTCTGGAAAGTGCAGGGTTGACCGTGTAGTTTGCTCCTCTCGGGGGCCTGTGGTGGCCATGGGGCAGGCGGCCTGG
GAGAGCTGCCGTACACAGCCACTGGGTGAGCCACACTCACGGTGGTAGAGCCACAGTGCCTGGTGCCACATCACGTCCT
CTGGATTTTAAAGTAAACACACACCTCCCGGCAGGCATCTGCCTGCGACCTGTGTGTGCCTGGGGAGAGTGGTAGCAC
GGAGGAAATTCGTGCACACTCAAGGTATCAGCAAGGTATCCGCAGTCAGGTGGAACGTGGAGGCCCTCTCTCTGGGATC
20 GTCTCCAGCGGATAAAGGACTGTGCACAGCTTCGGAAGCTTTATTTAAAAATATAACTATTAATTATTGCATTATAAGT
AATCACTAATGGTATCAGCAATTATAATATTTATTAAGTATAATAGAAATATTAAGTAGTACACACGTTCTGGAAAAA
CACAAATTGCACATGGCAGCAGAGTGAATTTTGGCCGAGGGACACGTGTGCACATGTGTGTAAGCGGCCCCAGGCCAC
AGAATTCGCTGACAAAGTCACTCCCCAGAGAAGCCACCACGGGCCTCCTTCGTGGTGTGAATTTTATTAAGATGGATC
AAGTCACGTACCGTCCACGTGTGGCAGGGCTTTGGGGAATGTGAGGTGATGACTGCGTCCTCATGCCCTGACAGACAGGA
25 GGTGACTGTGTCTGTCTGTCCCTAGGACACGGACAGGCCCGAAGCTCTAGTCCCCATCGTGGTCCAGTTTGGCCTCTGA
ATAAAAAACGCTCTCAAAACCTGTTGCCCCAAAACTAAGAACAGAGAGAGTTTCCCATCCCATGTGCTCACAGGGGCGTA
TCTGCTTGGCTTGAATCGCTGGGCTGGCCGGACTCCTAGAGTTGGTGCCTGTGCTTCTGTGCAAAAAGTGCAGTCCTCTT
GCCCATCACTGTGATATCTGCACCAGCAAGGAAAGCCTCTTTTCTTTTCTTTTCTTTTCTTTTGGAGACGGAACGTCA
CTGTTGTCTGCCTGGGCTTGAGTGCAGTGGCGGATCTCAACTCACTGCAACCTCCGCCTCCCGGGTCCAGCATTCTC
30 CTGCCTCAGCCTCCCGAGCAGCTGAGATTACAGGCACCCACCCCTGCGCCTGGCTAATTTTTGTATTTTGTAGTAGAG
GGGTTTTTGGCATGTTGGCCAGGCTGGTCTCGAATCCTGACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGTG
GGATTACAGGTGTGAGCCATCACGCCAGCCGAAAGCCTCTTTTAAAGGTGACCACCTATAGCGCTTCCCGAAAATAAC
AGGTCTTGTTTTTGCAGTAGGCTGCAAGCGTCTCTTAGCAACAGGAGTGGCGTCTGTGGGCTCTGGGGATGGCTGAGGG
TCGCGTGGCAGCCATGCCTTCTGTGTGCACCTTTAGGTCCACGGGGCTATTCTGCTCTCACTGTTTGTCTGAAAACGCA
35 CCTTGGCATCCTTGTGTTGGAGAGTTTCTGCTTCTCGTTGGTCTGCTGAAACTAGGGGCAAGGTTGTATCCGTTGGCGC
GCAGCGGCTACATGTAGGCTCATGAGTCTTTCACCGTGGACAAATTCCTTGAAAAAAGGAGTCCGGTTAAGCAT
TCATTCCGGGTCAAGTGTCTGTTCTGTGAATAAACTCTAAGATTTAAGAAACCTTAATGAAAGAAAACCTTGATGATTC
AGAGCAAGGATGTGGTACACCTGTGGCTGGATCTGTTTCAGCCGCCCCAGTGCATGGTGAGAGTGGGGAGCAGGGATTG
TTTGTTCAGAGGTCTCATCTGGTATGTTTCTGAGGTGTTTGCCGGCTGAATGGTAGACGTGTCGTTTGTGTGTATGAGGT
40 TCTGTGTCTGTGTGGCTCGGTTTGAGTGTACGCATGTCCAGCACATGCCCTGCCCGTCTCTCACCTGTGTCTTCCCGC
CCCAG

Intron 9 (SEQ ID NO 13)

45 GTGAGGCCTCCTCTTCCCAGGGGGCTTGGGTGGGGTTGATTTGCTTTTGATGCATTCACTGTTAATATTCCTGGTGC
TCTGGAGACCATGACTGCTCTGTCTTGAGGAACAGACAAGGTTGCAGCCCCCTCTTGGTATGAAGCCGCACGGGAGGGG

TTGCACAGCCTGAGGACTGCGGGCTCCACGCAGGCTCTGTCCAGCGGCCATGTCCAGAGGCCTCAGGGCTCAGCAGGCGG
GAGGGCCGCTGCCCTGCATGATGAGCATGTGAATCAACACCGAGGAAGCACACCAGCTTCTGTACGTCACCCAGGTTC
CGTTAGGGTCCCTTGGGGAGATGGGGCTGGTGCAGCCTGAGGCCCCACATCTCCAGCAGGCCCTCGACAGGTGGCCTGGA
CTGGGCGCCTCTTCAGCCCATTGCCCATCCCACTTGCATGGGGTCTACACCCAAGGACGCACACCTAAATATCGTGCC
5 AACCTAATGTGGTTCAACTCAGCTGGCTTTTATTGACAGCAGTTACTTTTTTTTTTTAATACTTTAAGTTC TAGGGTAC
ATGTGCACGACGTGCAGGTTAGTTACATATGTATACATGTGCCATGTTGGTGTGCTGCACCCATTAACATCATCATTACA
TTAGGTATATCTCCTAATGCTATCCCTCCCCACTCCCCCATCCCATGACAGGCCCTGGTGTGTGATGTTCCCCACCTG
TGTCCAAGTGTCTCATTGTTTCTCAGTTCACCTGTGAGTGAGAATGTGGTGTGTTGGTCTTTCTTCTTGCAATAGTTT
GCTCAGAGTGATGGTTTCCAGCTTCGTCCATGTCCCTACAAAGGACATGAACTCATCTTTTTTATGACTGCATAGTATT
10 CCGTGGTGTATATGTGCCACATTTTCTTAATCCAGTCTATCATCGATGGACATTTGGGTGGTTGCAAGTCTTTGCTACT
GTGAATAGTGCCGCAATAAACATACGTGTGCATGTGTCTTTATAGCAGCATGATTTATAATCCTTTGGGTATATACCCAG
TAATGGGATGGCTGGGTCAAATGGTATTTCTAGTTCTAGATCCTTGAGGAATCACCACACTGTCTCCACAATGGTTGAA
CTAGTTTACACTCCCACCAACAGTGTAAGAGTGTCTGGTGTGAGAGGATGTGGACAGCAGTTATTTTTTTATGAAAA
TAGTATCACTGAACAAGCAGACAGTTAGTGAAGGATGCGTCAGGAAGCCTGCAGGCCACACAGCCATTTCTCTCGAAGAC
15 TCCGGGTTTTCTGTGCATCTTTTGAACTCTAGCTCCAATTATAGCATGTACAGTGGATCAAGGTTCTTCTTCATTAA
GGTTCAAGTTCTAGATTGAAATAAGTTTATGTAACAGAAACAAAATTTCTGTACACACAACCTGCTCTGGGATTGGA
GGAAAGTGCTCGAGCTGGCGGCACACTGGTCAAGCCTCTGGGACAGGATACCTCTGGCCCATGGTCTGAGGGCGCTGG
GCTTGGGCTGAGGGTCACACAGTGCACCATGCCAGCTTCTGTGGATAGGATCTGGGTCTCGGATCATGCTGAGGACC
ACAGCTGCCATGCTGGTAAAGGGCACCACGTGGCTCAGAGGGGGCGAGGTTCCAGCCCCAGCTTTCTTACCGTCTTCAG
20 TTATTTTTCCCTAAGAGTCTGAGAAGTGGGGCCGCGCTGATGGCCTTCGTTCTGCTTCAGCTGGCACAGAAATGCACAA
GCTGATGGTAAACACTGAGTACTTATAATGAATGAGGAATTGCTGTAGCAGTTAACTGTAGAGAGCTCGTCTGTTGAAA
GAAATTTAAGTTTTTCATTTAACCCTTTGGAGAATGTTACTTTATTTATGGCTGTGTAATTTGTTTGACATTCAGTCCC
TCGTAGACAGATACTACGTAAAAAGTGTAAGTTAACCTTGCTGTGTATTTTCCCTATTTTAG

25 **Intron 10 (SEQ ID NO 14)**

GTGAGGCCCGTGCCGTGTGTCTGTGGGGACCTCCACAGCCTGTGGGCTTTGCAGTTGAGCCCCCGTGTCTGCCCTGG
CACCGCAGCGTGTGTCTCTGCCAAGTCCTCTCTCTCTGCCGCTGCTGGATCCGCAAGAGCAGAGGCGCTTGGCCGTGCACC
CAGGCCTGGGGGCGCAGGGGCACCTTCGGGAGGGAGTGGGTACCGTGCAGGCCCTGGTCTGCAGAGACGCACCCAGGTT
ACACACGTGGTGAGTGCAGGCGGTGACCTGGCTCCTGCTGCTCTTTGGAAAGTCAAGAGTGGCGGCTCCTGGGGCCCCAG
30 TGAGACCCCCAGGAGCTGTGCACAGGGCCTGCAGGGCCGAGGCGGCAGCCTCCTCCCCAGGGTGCACCTGAGCCTGCGGA
GAGCAGGAGCTGCTGAGTGAGCTGGCCACAGCGTTCGCTGCGGTACGTTCTGCTGGGGTTGTTTGGGATCGGTGGG
AGAATTTGGATTTGCTGAGTGCTGCTGTCTTGAACCACGGAGATGGCTAGGAGTGGGTTTCAGAGTTGATTTTGTGAAT
CAAATAAAATCAGGCACAGGGGACCTGGCCTCAGCACAGGGGATTGTCCAATGTGGTCCCCCTCAAGGGCGCCCCACAG
AGCCGGTGGGCTTGTTTTAAAGTGCATTTTACGAGGGACGAGAAACCTTGAAAGCTGTAAAGGGAACCTCAGAAAAATG
35 TGGCCGCCAGGGGTGGTTTCAGGTGCTTTGCTGGGCTGTGTTTGTGAAAACCCATTTGGACCCGCCCTCCAAGTCCACCC
TCCAGGTCCACCTCCAGGGCCGCCCTGGGCTGGGGGTATGCCTGGCGTTCTTGTGCCGACCCCGAGCACAGCAGGC
TGTGCACATTTAAATCCACTAAGATTCACTCGGGGGAGCCAGGTCCCAAGCAACTGAGGGCTCAGGAGTCTGAGGCT
GCTGAGGGGACAGAGCAGACGGGAACGCTGCTTCTGTGTGGCAAGTTCTGAGGGTGTGGCCAGGGAGGTGGCTCAGA
GTGTATGTTGGGGTCCCACCGGGGCGAGAACTCTGCTCTGATGAGTCGGCAGCCATGTAACAGGAAGGGGTGGCCACAG
40 GGAGCTGGGAATGCACCAGGGGAGCTGCGCAGCTGGCCGAGGTCCCAGGGCCAGGCCACAGGAAGGGCAGGGGGACGCC
GGGGCCACAGCAGAGGCCGAGGAAGGGAAGGGGATGCCAGGCCAGAGCAGAGGCTACCGGGCAGGGGGGCTCCCTG
AGCTGGGTGAGCGAGGCTCATGACTCGGCGAGGGAACCTCCTTGACGTGAAGCTGACGACTGGTGTGCCCCAGCTCACAG
CCCAGCCAGGTCCCGCGCTGAGCAGGAACCTCAGAACCCTCCCCTTTGTCTAAAGCACAGCAGATGCCTTCAGGGCATCT
AGGAGAAAAACAGGCAAAGTCGTTGAGAAACGCTCTAAAGAAGGTGGGATGGTGGCAATTTCTGTCCAGATTTAGTCT
45 GCCCGGACCACAGATGAGTCTATAACGGGATGTGGTGTGGCATGGGGACACATGAGATGGACCATCACAGAGGCCAC
TGGGGCTGCACCTCCCATCTGAGTCTGGCTGTCCCGGGTCCAGGCCAGGTTCTTGCACTGCTCACCTACCTGCTCCTGCC

GGGAGACAGGGAAAGCACCCCGAAGTCTGGAGCAGGGCTGGGTCCAGGCTCCTCAGAGCTCCTGCCAGGCCCAGCACCCCT
GCTCCAAATCACCCTTCTCTGGGGTTTTCCAAAGCATTTAACAAGGGTGTAGGTTACCTCCTGGGTGACGGCCCCGCA
TCCTGGGGCTGACATTGCCCTCTGCCTTAG

5 **Intron 11 (SEQ ID NO 15)**

GTGAGCGACCTGGCCGGAAGTGGAGCCTGTGCCCGGCTGGGGCAGGTGCTGCTGCAGGGCCGTTGCGTCCACCTCTGCT
TCCGTGTGGGGCAGGCGACTGCCAATCCCAAAGGGTCAAGGGCCACAGGGTGCCCTCGTCCCCTCTGGGGCTGAGCAGA
AATGCATCTTTCTGTGGGAGTGAGGGTGCTCACAACGGGAGCAGTCTTTCTGTGCTATTTTGGTAAAAGGAAATGGTGCAC
CAGACCTGGGTGCACTGAGGTGTCTTCAGAAAGCAGTCTGGATCCGAACCCAGACGCCCGGGCCCTGCTGGGCGTGAGT
10 CTCTCAAACCCGAACACAGGGGCCCTGCTGGGCATGAGTCCCTCTGAACCCGAGACCCCTGGGGCCCTGCTGGGCGTGAGT
CTCTCCGAACCCAGAGACTTCAGGGCCCTTTTGGGCGTGAGTCTCTCCGCTGTGAGCCCCACACTCCAAGGCTCATCCAC
AGTCTACAGGATGCCATGAGTTCATGATCACGTGTGACCCATCAGGGGACAGGGCCATGGTGTGGGGGGGTCTCTACAA
AATTCTGGGGTCTTGTTCCTCCAGAGCCCGAGAGCTCAAGGCCCCGTCTCAGGCTCAGACACAAATGAATTGAAGATGGA
CACAGATGCAGAAATCTGTGCTGTTTCTTTATGAATAAAAAGTATCAACATTCCAGGCAGGGCAAGGTGGCTCACACCT
15 ATAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCATTGAGGCCAGGAGTTTGAAGCCAACCTAACCAACATAGTG
AAATTCATTTCTACTTAAAAAATACAAAAATAGCCTGGCCTGGTGGCACACGCCTGTAGTCCCCGTATGCGGGAGGC
TGAGGCAGGAGAATCATTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCGAGATCACACCACTGCACTCCAGCCTGGGCA
ACAGAGTGAGACTTCATCTTAAAAAAGTATCAGCATTCCAAAACCATAGTGACAGGTGTTTTTTTATTC
TGTCCTTCGATAATATTTACTGGTGTGTGTAGAGGCCGGAAGTGGGGGTGCCCTCCTCTGAAAGGCACACCTTCATGG
20 GAAGAGAAATAAGTGGTGAATGGTTGTTAAACCAGAGGTTTAACTGGGGTCTGTGCTTCTGAGTTAACAGTCCAGATC
TGGACTTTGCCTCTTTCCAGAATGCTCCCTGGGGTTTGTCTCATGGGGGAGCAGCAGGTGTGGACACCCTCGTGATGGGG
GAGCAGCAGGTGCAGACGCCCTCATGATGGGGGAGTGGCAGGTGCAGACACCCTGTGCATGGTGGCCAGCATGTCCCTG
TTGCAGCTCCCTCCCAAGGATGCCGGTCTCCTGTGCTCCCCACAGTCCCTGCTTCCCTCTCACAGCCTTACCTGGTC
CTGGCCTCCACTGGCTTTGTCTGCATGATTTCCACATTTCTGGGCTCCAGCACCTCTTCGCTCTCCAGGCACCTCT
25 GCAGTGCTGGCCATACCAGTCAGCTGTGAAGTGTCCACTGCTTATTTTGTCTCCCATGAAATGTATTTTTTAGGACAGGC
ACCCCTGGTTCCAGCCTCTGGCAGCAGCATCAGTGAATGTTATTTGAAGGACAAAGGACAGACAAACAAATCAGGAAAATGG
GTTCTCTCTAAACACATTGCAAAGCCACAGAGGCTAGTGCAGGATGGGTGGGCATCAGGTGCATCAGATGTGGGTCCAATG
CCAGAATATTCTGTGCTCCCAAAGGCCACTTGGTCAGAGTGTGTGCTTGCAGAGGTGGCTCTAAAAGCTCAGCAGTGGAG
GCAGTGGTTCGCCATACTCAGGGTGAAGTCCATCCTCTGTGTCTGAAGTATACAGCAGAGGCTTGAAGGGCATCTGGGA
30 GAAGAAAACAGGCAAAATGATTAAGAAAAGTAAAAAGGAAAAGTGGTAAGATGGGAATTTTCTTGTCCAGATTTTAGTC
TCCCAAACCCAGCTCAGATGGTAGAATGTGGTCAGAACTGATGGACAGAACAAATAGAACAAAACGGAAGCCCTATCTCT
CAGAAAACGTGTGTTAATGTGGTATGTGGCACAGCTGATGGAAAAGAGAGTGTGTGTGTAATTTTTTTCTGAGAAAAC
GACTGGAAGCAAATAAGTTGTGTCTTTACAGCATATACCAGAGCAGATTCTAGGTAGAAGAGGAGACACATGCAACAAC
ACCAGCAACAGAAATAAAACAAAAGACTCAAAGGGAAGGGAGGTGAACGTTCCCTGGTTTGGTGTGGGGAAGGACACAC
35 AGGGAGGCGGATGAAACCACTGAGGCAACGGGCATTGCTTTCACTGCAGAGAACTCAGCTTGCTGAGCCACAGTGAAA
ATGGCCATTCCCTGGAGCGTTTGTGCACGTGATTTATTTAAGGCGCCCTGTGAGGTCTGCACATTATCCTCTCACTTT
GTTCTCCTAACCACCTGAGAGGTAGAGGAGGAAAGGCTCCAGGGGAGCAGCCGCCCTTGGTCAACCAGCTGGCAAAGGGC
ATGCATGATTGCAGCCTGGCCTCCTGCTCCGGGGCCCTTGTCTGCCCCAGGACCCACACAAGTCAGACCCATAGGCTC
AGGGTGAGCCGGAGCCCAAGGTCGTGTTGGGGATGGCTGTGAAAGAAGAAATGGACGCTCTGATGCACACTTGGGAAGGTC
40 CTACCAGCAGCGTCAAAGAAATGCATGTGAAACTGACAGCGAGACCCATCCCTCAAAGAAACGCACGTGAAACTGATGGC
GAGACCTGTCCCCATCCCTCATGCTGGCTCCTTTTCTGGGCTTGCCAAGAGCCAGCATCAGGTTGAGGCAAGCTGGAAAG
ACTTTTCTGGAAAGCAGCTTGTGTTGCATGGAAGTCTCACAATGTCTGTGTCTTCCAGTAATTCACCTTCTGAAGTGA
CCAGACATTATCAGGGTCTTATTTACCATTTCAGTGTTCAGGCAGGGGACTTGCCACAGCAAGTCACGAACCTGCC
CAAATACAGGGCTAAGGAGATATTATGCATCACAAAACCTTGCTCTGCCATTAAACATTTTCAAAGAATTTTGAAGAAT
45 GTTTAATGGCACAAAACGTTTATTTCAATGTAGCAGTGTTCAAAGCTGGATGTAAAAGAACACACCCCAAGGAGCCTGCCG
TGAATGTCTATGTGTGTTTCATCTTTGGACATGGACATACATGGGCAGTGAGTGGTGGTGAGGCCCCTGGAGGACATCGGTGG

001260-9122560

5

10

15

25

35

40

45

CCGGGACCTTAGGCTTATTTATTTGTTTAAAAACATTCTGGGCCTGGCTTCCGTTGTTGCTAAATGGGGAAAAGACATCC
CACCTCAGCAGAGTTACTGAGAGGCTGAAACCGGGGTGCTGGCTTGACTGGTGTGATCTCAGGTCATTCCAGAAGTGGCT
CAGGAAGTCAGTGAGACCAGGTACATGGGGGGCTCAGGCAGTGGGTGAGATGAGGTACACGGGGGGCTCAGGCAGTGGGT
GAGGCCAGGTACATGGGGGGCTCAGGCAGTGGGTGAGATGAGGTACACGGGGGGCTCAGGCAGAGGGTCAGACCAGGTAC
5 ACGGGGGCTCTGATCACACGCACATATGAGCACATGTGCACATGTGCTGTTTCATGGTAGCCAGGTCTGTGCACACCTGC
CCCAAAGTCCCAGGAAGCTGAGAGGCCAAAGATGGAGGCTGACAGGGCTGGCGCGGTGGCTCACACCTGTAGTCCCAGCA
CTTTGGGAGGCCGAGGCGAGAGGATCCCTTGAGCCCAGGAGTTTAAAGACCAGCCTGAGCAACATAGTAGAACCCCATCTC
TATGAAAAATAAAACAAAAATTAGCTGAACATGGTGGTGTGCGCCTGTAGTTCCAATACTTGGGAGGCTGAAGTGGGAG
GATCACTTGAGCCAGGAGGTGGAAGCTGCAGTGAGCTGAGATTGCACCACTGTACTGCAGCCTGGGTGACAGAGTGAGA
10 GCCCATCTCAACAACAACAAAGAAGACTGACAAATGCAGTTTCTTGAAAGAAACATTTAGTAGGAACCTAACCTACACA
CAGAAGCCAAGTCGGTGTCTCGGTGTGAGTGAGATGATGGGTCCTCACACCATCACCCAGACCCAGGGTTTATG
CACCACAGGGGCGGGTGGCTCAGAAGGGATGCGCAGGACGTTGATATACGATGACATCAAGGTTGTCTGACGAAGGGCAG
GATTCATGATAAGTACCTGCTGGTACACAAGGAACAATGGATAAACTGGAAACCTTAGAGGCCTTCCCGAACAGGGGCT
AATCAGAAGCCAGCATGGGGGGCTGGCATCCAGGATGGAGCTGCTTCAGCCTCCACATGCGTGTTCATACAGATGGTGCA
15 CAGAAACGCAGTGACCTGTGCACACACAGACACGCAGCTACTCGCACACACAAGCACACACAGACATGCATGCATGC
ATCCGTGTGTGTGCACCTGTGCCCATGAGGAAACCCATGCATGTGCATTTCATGCACGCACACAGGCACCGGTGGGCCCAT
GCCCACACCCACGAGCACCCTCTGATTAGGAGGCCTTCTCTGACGCTGTCCGCCATCCTCTCAG

Intron 14 (WEQ ID NO 18)

GTATGTGCAGGTGCCTGGCCTCAGTGGCAGCAGTGCCTGCCTGCTGGTGTAGTGTGTCAGGAGACTGAGTGAATCTGGG
CTTAGGAAGTTCTTACCCCTTTTCGCATCAGGAAGTGGTTTAAACCAACCACTGTCAGGCTCGTCTGCCCCCCTCTCGT
GGGGTGAGCAGAGCACCTGATGGAAGGGACAGGAGCTGTCTGGGAGCTGCCATCCTTCCACCTTGCTCTGCCTGGGGAA
GCGCTGGGGGGCCTGGTCTCTCCTGTTTGCCCCATGGTGGGATTGGGGGGCCTGGCCTCTCCTGTTTGCCCTGTGGTGG
GATTGGGCTGTCTCCCGTCCATGGCACTTAGGGCCCTTGTGCAAACCCAGGCCAAGGGCTTAGGAGGAGGCCAGGCCAG
25 GCTACCCACCCCTCTCAGGAGCAGAGGCCGCTATCACCACGACAGAGCCCCGCGCGTCTCTGCTTCCAGTCACCG
TCCTCTGCCCCCTGGACACTTTGTCCAGCATCAGGGAGGTTTCTGATCCGTCTGAAATTCAAGCCATGTCGAACCTGCGGT
CCTGAGCTTAACAGCTTCTACTTTCTGTTCTTCTGTGTTGTGGAAATTTACCTGGAGAAGCCGAAGAAACATTTCTG
TCGTGACTCCTGCGGTGCTTGGGTGCGGACAGCCAGAGATGGAGCCACCCCGCAGACCGTGGGGTGTGGGCAGCTTTCCG
GTGTCTCCTGGGAGGGGAGCTGGGCTGGGCCTGTGACTCCTCAGCCTCTGTTTTCCCCCAG

Intron 15 (SEQ ID NO 19)

GCAAGTGTGGGTGGAGGCCAGTGCGGGCCCCACCTGCCCAGGGGTATCCTTGAACGCCCTGTGTGGGGCGAGCAGCCTC
AGATGCTGCTGAAGTGCAGACGCCCCCGGGCCTGACCCTGGGGGCTGGAGCCACGCTGGCAGCCCTATGTGATTAAACG
CTGGTGTCCCCAGGCCACGGAGCCTGGCAGGGTCCCCAATTCTTGAACCCCTGCTTCCCATCTCAGGGGCGATGGCTCC
35 CCACGCTTGGGAGCCTTCTGACCCCTGACCTGTGTCTCTCACAGCCTCTTCCCTGGCTGCTGCCCTGAGCTCCTGGGGT
CCTGAGCAAGTTCTCTCCCCGCCCCGCGCTCCAGCGTCACTGGGTGCTGTCTGCTCGCCCCGGTGGAGGGGTGTCTG
TCCCTTCACTGAGGTTCCCACCAGCCAGGGCCACGAGGTGCAGGCCCTGCCTGCCCGGCCACCCACACGTCCTAGGAGGG
TTGGAGGATGCCACCTCTGGCCTCTTCTGGAACGGAGTCTGATTTTGGCCCCGAG

3'-untranscribed region (SEQ ID NO 20)

ATCTCATGTTTGAATCCTAATGTGCACTGCATAGACACCACTGTATGCAATTACAGAAGCCTGTGAGTGAACGGGGTGGT
GGTCAGTGCGGGGCCCATGGCCTGGCTGTGCATTTACGGAAGTCTATGAGTGAATGGGGTTGTGGTCAGTGGGGGCCCATG
GCCTGGCTGGGCGCTGGGAGGTTTCTGATGCTGTGAGGCAGGAGGGGAAGGAGGGTAGGGGATAGACAGTGGGAGCCCCCA
CCCTGGAAGACATAACAGTAAGTCCAGGCCCGAAGGGCAGCAGGGATGCTGGGGGCCAGCTTGGGCGGCGGGGATGATG
45 GAGGGCCTGGCCAGGGTGGCAGGGATGATGGGGGCCCGAGCTGGGGTGGCAGGGGTGATGGGGGGGGCTGGTCTGGGTGG

004260:42350

CGGGGAAGATGGGGAAGCCTGGCTGGGCCCCCTCCTCCCCTGCCTCCACCTGCAGCCGTGGATCCGGATGTGCTTCCCT
GGTGCACATCCTCTGGGCCATCAGCTTTCATGGAGGTGGGGGAGGGGCATGACACCATCCTGTATAAAATCCAGGATT
CCTCCTCCTGAACGCCCCAACTCAGGTTGAAAGTCACATTCCGCCCTCTGGCCATTCTCTTAAGAGTAGACCAGGATTCTG
ATCTCTGAAGGGTGGGTAGGGTGGGCGAGTGGAGGGTGTGGACACAGGAGGCTTCAGGGTGGGGCTGGTGATGCTCTCTC
5 ATCCTCTTATCATCTCCCAGTCTCATCTCTCATCTCTTATCATCTCCCAGTCTCATCTGTCTTCTCTTATCTCCCAGT
CTCATCTGTCATCCTCTTACCATCTCCCAGTCTCATCTCTTATCCTCTTATCTCCTAGTCTCATCCAGACTTACCTCCCA
GGGCGGGTGCCAGGCTCGCAGTGGAGCTGGACATACGTCCTTCTCAGGCAGAAGGAACTGGAAGGATTGCAGAGAACAG
GAGGGGCGGCTCAGAGGGACGCAGTCTTGGGGTGAAGAAACAGCCCCCTCCTCAGAAGTTGGCTTGGGCCACACGAAACCG
AGGGCCCTGCGTGAGTGGCTCCAGAGCCTTCCAGCAGGTCCCTGGTGGGGCCTTATGGTATGGCCGGGTCTACTGAGTG
10 CACCTTGGACAGGGCTTCTGGTTTGAAGTGCAGCCCGACGTGCCTGGTGTGGGGTGGGGGCTTATGGCCACTGGATATG
GCGTCATTTATTGCTGCTGCTTCAGAGAATGTCTGAGTGACCGAGCCTAATGTGTATGGTGGGCCCCAAGTCCACAGACTG
TGTCGTAATGCACTCTGGTGCCTGGAGCCCCCGTATAGGAGCTGTGAGGAAGGAGGGGCTCTTGGCAGCCGGCCTGGGG
GCGCCTTTGCCCTGCAAACTGGAAGGGAGCGCCCCGGGCGCCGTGGGCGGACGACCTCAAGTGAGAGGTTGGACAGAAC
AGGGCGGGGACTTCCCAGGAGCAGAGGCCGCTGCTCAGGCACACCTGGGTTTGAATCACAGACCAACaGGTCAGGCCATT
15 GTTCAGCTATCCATCTTCTACAAAGCTCCAGATTCCTGTTTCTCCGGGTGTTTTTTTGTGAAATTTTACTCAGGATTACT
TATATTTTTTGCTAAAGTATTAGACCCTTAAAAAAGGTATTTGCTTTGATATGGCTTAACTCACTAAGCACCTACTTTAT
TTGTCTGTTTTTATTTATTATTATTATTATTATTAGAGATGGTGTCTACTCTGTCACCCAGGTTGTTAGTGCAGTGGCAC
AGTCATGGCTCGCTGTAGCCGCAAACCCCAGGCTCAAGTGATCCTCCGGCCTCAGCTTCCCAGAGTGTGGGATTACAG
GTGTGAGCCACTGCCCTTGCTGGCACTTTTAAAAACCCTATGTAAGGTGAGTCCAGTGGCTTCCACACCTGTCATCC
20 CAGTAGTTTGGGAAGCCGAGGCAGAAGGATTGTCTGAGGCCAGGAGTTTGAAGCCAGCATGGGTAACATAGGGAGACCCC
ATCTCTACAAAAATGCAAAAAGTTATCCGGGCGTGGGGTCCAGCATCTGTAGTCCCAGCTGCTCGGGAGGCTGAGTGGG
AGGATCGCTTGAGCCCGGAGGTGATGGCTGCAGTGAGCTGTGATTGTACCATCGCACTCCAGCCTGGGCAACAGAGTGA
GACCTGTCTCAAAAAAAAAAAAAAAAAAAGGAGAAGGAGAAGAGAAGAAGGAAGAAGGAAGAAGAGAAGAAG
GAAGAAGGAAGAAAGAAGGAGAAGGAGGCCCTGCTAGGTGCTAGGTAGACTGTCAAATCTCAGAGCAAAATGAAAATAACA
25 AAGTTTTAAAGGGAAGAAAAACCCAGCTCTTTGGACTTCCTTAGGCCTGAACCTCATCTCAAGCAGCTTCCTTCCACA
GACAAGCGTGATGGAGCGAGTGAGTTCAAAGCAGAAAGGGAGGAGAAGCAGGCAAGGGTGGAGGCCTGTGGGTGACACCA
GCCAGGACCCCTGAAAGGGAGTGGTTGTTTTCTGCCTCAGCCCCACGCTCCTGCCGGTCTGCACCTGCTGTAACCGTC
GATGTTGGTGCCAGGTGCCACCTGGGAAGGATGCTGTGCAGGGGGCTTGCCAACTTTGGTGGGTTTCAGAAGCCCCAG
GCACCTTGTTGGCAGGCACAATTACAGCCCCCTCCCCAAAGATGCCACGTCCTTCTCCTGGAACCTGTGAATGTGTACCCG
30 CAAGGCAGAGGCTGGTGAAGGCTGCAGGTGGAATCACGGCTGCCAGTCAGCCGATCTTAAGGTCATCCTGGATTATCTGG
TGGGCCTGATATGGCCACAAGGGTCCCTAGAAGTGAGAGAGGGAGGCAGGGGAGAGTCAGAGAGGGGACGTGAGAAGGAC
CACTGGCCACTGCTGGCTTTGAGATGGAGGAGGGGGTCCCCAGCCAAGGAATGGGGGCAGCCGCTCCATGCTGGAAAAGC
AAGCAATCCTCCCCGGTCTGAGGGCACACGGCCCTGCCACGCTCGATTTCAGGCCAGTGGGACCTGTTTCAGCTTTC
CGGCCTCCAGAGCTGTAAGATGATGCGTTTGTGTTAGCCACTAAGCTGCAGTGATTGTCACAGCAGCAATGGAATAG
35 CAGTACAGGGAAATGAATACAGGGACAGTTCTCAGAGTGAATCTCAGCCACCCCTGGG

00160" 912360

Characterization of the exons showed, interestingly, that the functionally important hTC protein domains which are described in our Patent Application PCT/EP/98/03469 are arranged on separate exons. The telomerase-characteristic T motif is located on exon 3. The RT (reverse transcriptase) motifs 1-7, which are important for the catalytic function of the telomerase, are located on the following exons: RT motifs 1 and 2 on exon 4, RT motif 4 on exon 9, RT motif 5 on exon 10, and RT motifs 6 and 7 on exon 11. RT motif 3 is shared by exons 5 and 6 (see Fig. 8).

Elucidation of the exon-intron structure of the hTC gene also shows that the four deletions or insertion variants of the hTC cDNA which were described in our Patent Application PCT/EP/98/03469, as well as three additional hTC insertion variants which are described in the literature (Kilian et al., 1997), in all probability represent alternative splicing products. As shown in Fig. 8, the splicing variants can be divided into two groups: deletion variants and insertion variants.

The hTC variants in the deletion group lack specific sequence segments. The 36 bp in-frame deletion in variant DEL1 in all probability results from using an alternative 3' splice acceptor sequence in exon 6, resulting in a part of RT motif 3 being lost. In variant DEL2, the normal 5' splice donor and 3' splice acceptor sequences of introns 6, 7 and 8 are not used. Instead exon 6 is fused directly to exon 9, resulting in a displacement arising in the open reading frame and a stop codon appearing in exon 10. Variant Del3 is a combination of variants 1 and 2.

The insertion variant group is characterized by the insertion of intron sequences which lead to premature cessation of translation. Instead of the 5' splice donor sequence of intron 5, which is normally used, use is made, in variant INS1, of an alternative, 3'-located splice site, resulting in the insertion of the first 38 bp from intron 4 between exon 4 and exon 5. The insertion, in variant INS2, of a region of the intron 11 sequence likewise results from using an alternative 5' splice donor sequence in intron 11. Since this variant was only described inadequately in the

literature (Kilian et al., 1997), it is not possible to determine the precise alternative 5' splice donor sequence in this variant. The insertion of intron 14 sequences between exon 14 and exon 15 in variant INS3 comes from using an alternative 3' splice acceptor sequence, resulting in the 3' part of intron 14 not being spliced.

5

The hTC variant INS4 (variante 4), which is described in our Patent Application PCT/EP/98/03469, is characterized by exon 15, and the 5' part region of exon 16, being replaced by the first 600 bp of intron 14. This variant can be attributed to the use of an alternative internal 5' splice donor sequence in intron 14 and an alternative 3' splice acceptor sequence in exon 16, resulting in an altered C terminus.

10

The *in vivo* generation of hTC protein variants which are probably non-functional and which could interfere with the function of the complete hTC protein constitutes a possible mechanism, in addition to transcription regulation, for controlling hTC protein function. The function of the hTC splicing variants is not yet known. Although most of these variants presumably encode proteins without reverse transcriptase activity, they could nevertheless play a crucial role as transdominant-negative telomerase regulators by, for example, competing for interaction with important binding partners.

15

20

The search for possible transcription factor binding sites was carried out using the „find pattern“ algorithm from the Genetics Computer Group (Madison, USA) GCG Sequence Analysis program package. This resulted in the identification of a variety of potential binding sites for transcription factors in the nucleotide sequence of intron 2, which binding sites are listed in Tab. 2. In addition, an Sp1 binding site was found in intron 1 (pos. 43), and a c-Myc binding site was found in the 5'-untranslated region (cDNA position 29-34, cf. Fig. 6).

25

001260-34233560

Example 6

In order to ascertain the start point(s) of hTC transcription in HL 60 cells, the 5' end of the hTC mRNA was determined by means of primer extension analysis.

5

2 µg of polyA⁺ RNA from HL-60 cells were denaturated at 65°C for 10 min. 1 µl of RNasin (30-40 U/ml) and 0.3-1 pmol of radioactively labelled primer (5'GTTAAGTTGTAGCTTACACTGGTTCTC 3'; 2.5-8x10⁵ cpm) were added for primer annealing, and the whole was incubated, at 37°C for 30 min, in a total volume of 20 µl. After the addition of 10 µl of 5xreverse transcriptase buffer (from Gibco-BRL), 2 µl of 10 mM dNTPs, 2 µl RNasin (see above), 5 µl of 0.1 M DTT (from Gibco-BRL) 2 µl of ThermoScript RT (15 U/µl; from Gibco-BRL) and 9 µl of DEPC-treated water, primer extension took place, at 58°C for 1 h, in a total volume [lacuna]. The reaction was stopped by adding 4 µl of 0.5 M EDTA, pH 8.0, and the RNA was degraded, at 37°C for 30 min, after having added 1 µl of RNaseA (10 mg/ml). 2.5 µg of sheared calf thymus DNA and 100 µl of TE were then added, and the mixture was extracted once with 150 µl of phenol/chloroform (1:1). The DNA was precipitated, at -70°C for 45 min, after adding 15 µl of 3 M Na acetate and 450 µl of ethanol, and then centrifuged at 14,000 rpm for 15 min. The precipitate was washed once with 70% ethanol, dried in air and dissolved in 8 µl of sequencing stop solution. After 5 min of denaturation at 80°C, the samples were loaded onto a 6% polyacrylamide gel and fractionated electrophoretically (Ausubel et al., 1987) (Fig. 5).

25 In this connection, a main transcription start site was identified which is located 1767 bp 5' of the ATG start codon of the hTC cDNA sequence (nucleotide position 3346 in Fig. 4). In addition to this, the nucleotide sequence around this main transcription start (TTA₊₁TTGT) represents an initiator element (Inr), which, in 6 out of 7 nucleotides, matches the consensus motif (PyPyA₊₁Na/tPyPy) (Smale, 1997) of
30 an initiator element.

It was not possible to identify any unambiguous TATA box in the immediate vicinity of the experimentally identified main transcription start, which means that the hTC promoter has probably to be classified in the family of TATA-less promoters (Smale, 1997). However, a potential TATA box from nucleotide position 1306 to nucleotide position 1311 (Fig. 4) was found by means of bioinformatics analysis. The subsidiary transcription starts which were additionally observed around the main transcription start have also been described in the case of other TATA-less promoters (Geng and Johnson, 1993), for example in the strongly regulated promoters of some cell cycle genes (Wick *et al.*, 1995).

Example 7

In addition to the start point of the hTC transcript which was described in Example 6 and identified in HL60 cells, a further transcription start region was also identified in HL60 cells. With the aid of RT-PCR analyses, the region of the hTC gene transcription start in HL60 cells was localized to bp -60 to bp -105.

The cDNA for this was synthesized using a First Strand cDNA Synthesis kit (Clontech), in accordance with the manufacturer's instructions, and employing 0.4 µg of HL60 cell polyA RNA (Clontech) and the gene-specific primer GSP13 (5'-CCTCCAAAGAGGTGGCTTCTTCGGC-3', cDNA position 920-897). In a final volume of 50 µl, 10 pmol dNTP mix were added to 1 µl of cDNA, and a PCR reaction was carried out in 1xPCR reaction buffer F (PCR-Optimizer kit from InVitrogen) and using one unit of platinum Taq DNA polymerase (from Gibco/BRL). 10 pmol of each of the 5' and 3' primers defined below were added as primers. The PCR was carried out in 3 steps. A two-minute denaturation at 94°C was followed by 36 PCR cycles in which the DNA was first of all denatured at 94°C for 45 sec and, after that, the primers were annealed, and the DNA chain was extended at 68°C for 5 min. The cycles were concluded by a chain extension at 68°C for 10 min. In all, six different 5' PCR primers (primer HTRT5B: 5'-CGCAGCCACTACCGCGAGGTGC-3', cDNA position 105 to 126; primer C5S:

5'-CTGCGTCCTGCTGCGCACGTGGGAAGC-3', 5'-flanking region -49 to -23; primer PRO-TEST1: 5'-CTCGCGGCGCGAGTTTCAGGCAG-3', 5'-flanking region -74 to -52; primer PRO-TEST2: 5'-CCAGCCCCTCCCCTTCCTTTCC-3', 5'-flanking region -112 to -91; primer PRO-TEST4: 5'-CCAGCTCCGCCTCCTCCGCGC-3', 5'-flanking region -191 to -171; primer RP-3A: 5'-CTAGGCCGATTCGACCTCTCTCC-3', 5'-flanking region -427 to -405) were combined with the 3' PCR primer C5Rback (5'-GTCCCAGGGCACGCACACCAG-3', cDNA position 245 to 225). Genomic DNA was also employed for the PCR, as a control, in addition to the Oligo dT- and GSP13-primed cDNAs. As Fig. 9 shows, a PCR product was only obtained with the primer combinations HTRT5B-C5Rback, C5S-C5Rback and PRO-TEST1-C5Rback, indicating that the start point for hTC transcription lies in the region between bp-60 and bp-105.

Example 8

Several extremely GC-rich regions, so-called CpG Islands, are located in the isolated 5'-flanking region, of about 11.2 kb in size, of the hTC gene. One CpG Island, having a GC content of > 70%, extends from bp - 1214 into intron 2. Two further GC-rich regions having a GC content of > 60% extend from bp -3872 to bp -3113 and from bp -5363 to bp -3941, respectively. The positions of the CpG Islands are shown graphically in Fig. 11.

The search for possible transcription factor binding sites was carried out using the "Find Pattern" algorithm from the Genetics Computer Group (Madison, USA) GCG Sequence Analysis program package. This resulted in the identification of a variety of potential binding sites in the region up to -900 bp upstream of the translation start codon ATG: five Sp1 binding sites, one c-Myc binding site, and one CCAC box (Fig. 10). In addition, a CCAAT box and a second c-Myc binding site were found at positions -1788 and -3995, respectively, of the 5'-flanking region.

Example 9

In order to analyse the activity of the hTC promoter, PCR amplification was used to generate four hTC promoter sequence segments of differing length, which segments were cloned into the Promega vector pGL2 5' in front of the luciferase reporter gene. The 8.5 kb SacI fragment which was subcloned from phage clone P12 was selected as the DNA source for the PCR amplification. In a final volume of 50 µl, 10 pmol of dNTP mix were added to 35 ng of this DNA, and a PCR reaction was carried out in 1xPCR reaction buffer (PCR-Optimizer kit from InVitrogen) and using one unit of platinum Taq DNA polymerase (from Gibco/BRL). In each case 20 pmol of the 5' and 3' primers which are defined below were added as primers. The PCR was carried out in three steps. A two-minute denaturation at 94°C was followed by 30 PCR cycles in which the DNA was first of all denaturated at 94°C for 45 sec, after which the primers were annealed, and the DNA chain was extended, at 68°C for 5 min. The cycles were concluded by a chain extension at 68°C for 10 min. The selected 3' PCR primer was in each case the primer PK-3A (5'-GCAAGCTTGACGCAGCGCTGCCTGAACTCG-3', position -43 to -65), which primer recognizes a sequence region 42 bp upstream of the ATG START codon. A promoter fragment of 4051 bp in size (NPK8) was amplified by combining the PK-3A primers with the 5' PCR primer PK-5B (5'-CCAGATCTCTGGAACACAGAGTGGCAGTTTCC-3', position -4093 to -4070). Combining the pair of primers PK-3A and PK-5C (5'-CCAGATCTGCATGAAGTGTGTGGGGATTTGCAG-3', position -3120 to -3096) led to the amplification of a promoter fragment of 3078 bp in size (NPK15). Use of the primer combination PK-3A and PK-5D (5'-GGAGATCTGATCTTGGCTTACTGCAGCCTCTG-3', position -2110 to -2087) amplified a promoter fragment of 2068 bp in size (NPK22). Finally, using the primer combination PK-3A and PK-5E (5'-GGAGATCTGTCTGGATTCCTGGGAAGTCCTCA-3', position -1125 to -1102) led to the amplification of a promoter fragment of 1083 bp in size (NPK27).

The PK-3A primer contains a HindIII recognition sequence. The different 5' primers contain a BglII recognition sequence.

5 The resulting PCR products were purified using the Qiagen QIA quick spin PCR purification kit, in accordance with the manufacturer's instructions, and then digested with the restriction enzymes BglII and HindIII. The pGL2 promoter vector was digested with the same restriction enzymes, and the SV40 promoter contained in this vector was released and removed. The PCR promoter fragments ligated into the vector, which was then transformed into competent DH5 α bacteria (from
10 Gibco/BRL). DNA for the promoter activity analyses, which are described below, was isolated from transformed bacterial clones using the Qiagen plasmid kit.

Example 10

15 The activity of the hTC promoter was analysed in transient transfections in eukaryotic cells.

All the work with eukaryotic cells was carried out at a sterile workstation. CHO-K1 and HEK 293 cells were obtained from the American Type Culture collection.

20 CHO-K1 cells were kept in DMEM Nut Mix F-12 cell culture medium (from Gibco-BRL, order number: 21331-020) containing 0.15% streptomycin/penicillin, 2 mM glutamine and 10% FCS (from Gibco-BRL).

25 HEK 293 cells were cultured in DMOD cell culture medium (from Gibco-BRL, order number: 41965-039) containing 0.15% streptomycin/penicillin, 2 mM glutamine and 10% FCS (from Gibco-BRL).

30 CHO-K1 and HEK 293 cells were cultured at 37°C in a water-saturated atmosphere while being gassed with 5% CO₂. When the cell lawn was confluent, the medium was sucked off, after which the cells were washed with PBS (100 mM KH₂PO₄ pH

7.2; 150 mM NaCl) and released by adding a trypsin-EDTA solution (from Gibco-BRL). The trypsin was inactivated by adding medium and the cell count was determined using a Neubauer counting chamber in order to plate out the cells at the desired density.

5

For the transfection, in each case 2×10^5 HEK 293 cells were plated out, per well, in a 24-well cell culture plate. The HEK 293 medium was removed after 3 hours. For the transfection, up to 2.5 μ g of plasmid DNA, 1 μ g of a CMV β -Gal plasmid construct (from Stratagene, order number: 200388), 200 μ l of serum-free medium and 10 μ l of transfection reagent (DOTAP from Boehringer Mannheim) were incubated at room temperature for 15 minutes and then dropped uniformly onto the HEK 293 cells. 1.5 ml of medium were added after 3 hours. The medium was changed after 20 hours. After a further 24 hours, the cells were harvested for determining the luciferase activity and the β -Gal activity. For this, the cells were lysed, at room temperature for 15 minutes, in the cell culture lysis reagent (25 mM Tris [pH 7.8] containing H_3PO_4 ; 2 mM CDTA; 2 mM DTT; 10% glycerol; 1% Triton X-100). Twenty μ l of this cell lysate were mixed with 100 μ l of luciferase assay buffer (20 mM Tricin; 1.07 mM $(\text{MgCO}_3)_4$ $\text{Mg}(\text{OH})_2 \cdot 5\text{H}_2\text{O}$; 2.67 mM MgSO_4 ; 0.1 mM EDTA; 33.3 mM DTT; 270 μ M coenzyme A; 470 μ M luciferin, 530 μ M ATP), and the light generated by the luciferase was measured.

20

In order to measure the β -galactosidase activity, equal quantities of cell lysate and β -galactosidase assay buffer (100 mM sodium phosphate buffer, pH 7.3; 1 mM MgCl_2 ; 50 mM β -mercaptoethanol; 0.665 mg of ONPG/ml) were incubated at 37°C for at least 30 minutes or until a slight yellow coloration appeared. The reaction was stopped by adding 100 μ l of 1 M Na_2CO_3 , and the absorption was determined at 420 nm.

25

In order to analyse the hTC promoter, four hTC promoter sequence segments of differing length were cloned 5' in front of the luciferase reporter gene (cf. Example 9).

30

007250:Str20360

5

10

also observed in CHO cells (data not shown).

References

- 5 **Allsopp, R. C., Vazire, H., Pattersson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W. und Harley, C.B. (1992).** Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. **89**, 10114-10118.
- 10 **Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1987).** Current protocols in molecular biology. Greene Publishing Associates and Wiley-Intersciences, New York.
- 15 **Blasco, M. A., Rizen, M., Greider, C. W. und Hanahan, D. (1996).** Differential regulation of telomerase activity and telomerase RNA during multistage tumorigenesis. Nature Genetics **12**, 200-204.
- 20 **Broccoli, D., Young, J. W. und deLange, T. (1995).** Telomerase activity in normal and malignant hematopoietic cells. Proc. Natl. Acad. Sci. **92**, 9082-9086.
- 25 **Counter, C. M., Ailion, A. A., LeFeuvre, C. E., Stewart, N. G. Greider, C.W. Harley, C. B. und Bacchetti S. (1992).** Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. EMBO J. **11**, 1921-1929.
- 30 **Feng, J., Funk, W. D., Wang, S.-S., Weinrich, S. L., Ailion, A.A., Chiu, C.-P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., Le, S., West, M.D., Harley, C.B., Andrews, W.H., Greider, C.W. und Villeponteau, B. (1995).** The RNA component of human telomerase. Science **269**, 1236-1241.
- 35 **Geng, Y., and Johnson, L.F. (1993).** Lack of an initiator element is responsible for multiple transcriptional initiation sites of the TATA less mouse thymidine synthase promoter. Mol. Cell. Biol **14**:4894.
- Goldstein, S. (1990).** Replicative senescence: The human fibroblast comes of age. Science **249**, 1129-1133.
- Harley, C.B., Futcher, A.B., Greider, C.W., 1990.** Telomeres shorten during ageing of human fibroblasts. Nature **345**, 458-460.

- Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D.K. und Allshire, R.C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. *Nature* **346**, 866-868.
- 5 Hiyama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyama, E., Piatyszek, M.A., Shay, J.W., Ishioka, S. und Yamakido, M. (1995). Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. *J. Immunol.* **155**, 3711-3715.
- 10 Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C. B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L. und Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015.
- 15 Latchman, D.S. (1991). Eukaryotic transcription factors. Academic Press Limited, London.
- 20 Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. und Cech T.R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* **276**: 561-567.
- 25 Lundblad, V. und Szostak, J. W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* **57**, 633-643.
- 30 McClintock, B. (1941). The stability of broken ends of chromosomes in *Zea mays*. *Genetics* **26**, 234-282.
- 35 Meyne, J., Ratliff, R. L. und Moyzis, R. K. (1989). Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proc. Natl. Acad. Sci.* **86**, 7049-7053.
- Olovnikov, A. M. (1973). A theory of marginotomy. *J. Theor. Biol.* **41**, 181-190.
- Sandell, L. L. und Zakian, V. A. (1993). Loss of a yeast telomere: Arrest, recovery and chromosome loss. *Cell* **75**, 729-739.
- Shapiro, M.B., Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids Res.* **15**, 7155-7174.
- Smale, S.T. and Baltimore, D. (1989). The „initiator“ as a transcription control element. *Cell* **57**:103-113.

Smale, S.T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta* 1351, 73-88.

5 **Shay, J. W. (1997).** Telomerae and Cancer. Ciba Foundation Meeting: Telomeres and Telomerase. London.

10 **Vaziri, H., Dragowska, W., Allsopp, R. C., Thomas, T. E., Harley, C.B. und Landsdorp, P.M. (1994).** Evidence for a mitotic clock in human hematopoietic stem cells: Loss of telomeric DNA with age. *Proc. Natl. Acad. Sci.* **91**, 9857-9860.

15 **Wick, M., Härönen, R., Mumberg, D., Bürger, C., Olsen, B.R., Budarf, M.L., Apte, S. S. and Müller, R. (1995).** Structure of the human TIMP-3 gene and its cell-cycle-regulated promoter. *Biochemical Journal* 311, 549-554.

Zakian, V. A. (1995). Telomeres: Beginning to understand the end. *Science* **270**, 1601-1607.

001260: 9123360

SEQUENCE LISTING

<110> Bayer AG

5 <120> Regulatory DNA sequences from the 5i region of the gene
for the human catalytic telomerase subunit, and
their diagnostic and therapeutic use

10 <130> LeA32805-Foreign Countries

<140>
<141>

15 <160> 20

<170> PatentIn Vers. 2.0

<210> 1
<211> 5126
<212> DNA
20 <213> Homo sapiens

<400> 1

| | | | | | | | |
|----|-------------|-------------|------------|-------------|-------------|--------------|------|
| 25 | gagctctgaa | ccgtggaaac | gaacatgacc | cttgccctgcc | tgcttccctg | ggtaggggtcaa | 60 |
| | gggtaaatgaa | gtgggtgtgca | ggaaatggcc | atgtaaatta | cacgactctg | ctgatgggga | 120 |
| | ccgttccttc | catcattatt | catcttcacc | cccaaggact | gaatgattcc | agcaacttct | 180 |
| | tcgggtgtga | caagccatga | caaaactcag | tacaaacacc | actcttttac | taggcccaca | 240 |
| | gagcacgggc | cacacccctg | atatattaag | agtccaggag | agatgaggct | gctttcagcc | 300 |
| | accaggctgg | ggtagacaaca | gcggctgaac | agtctgttcc | tctagactag | tagaccctgg | 360 |
| 30 | caggcactcc | cccaaattct | agggcctggt | tgctgcttcc | cgagggcgcc | atctgccctg | 420 |
| | gagactcagc | ctgggggtgcc | acactgaggg | cagccctgtc | tccacaccct | ccgcctccag | 480 |
| | gcctcagctt | ctccagcagc | ttcctaaacc | ctgggtgggc | cggtgtccag | cgctactgtc | 540 |
| | tcacctgtcc | cactgtgtct | tgtctcagcg | acgtagctcg | cacggttcct | cctcacatgg | 600 |
| | gggtgtctgt | tccttcccca | acactcacat | gcgttgaagg | gaggagattc | tgcgccctcc | 660 |
| 35 | agactggctc | ctctgagcct | gaacctggct | cgtagggccc | gatgcagggt | cctggcgctc | 720 |
| | ggctgcacgc | tgacctccat | ttccaggcgc | tccccgtctc | ctgtcatctg | ccggggcctg | 780 |
| | ccgggtgtgt | cttctgtttc | tgtgtctcct | ttccacgtcca | gctgcgtgtg | tctctgcccg | 840 |
| | ctaggggtctc | gggggttttta | taggcatagg | acggggggcgt | ggtagggccag | ggcgctcttg | 900 |
| 40 | ggaaatgcaa | catttgggtg | tgaaagttag | agtgccctgtc | ctcacctagg | ttccacggga | 960 |
| | caggcctggg | gatggagccc | ccgccaggga | cccgcccttc | tctgcccagc | actttctctg | 1020 |
| | ccccctccct | ctggaacaca | gagtggcagt | ttccacaagc | actaagcctc | ctcttcccaa | 1080 |
| | aagaccagcg | attggcaccc | ctggacattt | gcccacagc | cctgggaatt | cagctgcagg | 1140 |
| | cgcacatcat | gtacacactc | ccgtccacga | ccgacccccc | ctgttttatt | ttaatagcta | 1200 |
| 45 | caaagcaggg | aaatccctgc | taaaatgtcc | tttaacaaac | tggttaaaca | aacgggtcca | 1260 |
| | ttccgcagcg | ggacagtctc | tcacagtga | gaggaacatg | ccgtttataa | ccgtgttctg | 1320 |
| | catctcaagg | gaattacgct | gagtcaaaac | tgccacctcc | atgggatacg | tacgcaacat | 1380 |
| | gctcaaaaag | aaagaatttc | accccatggc | aggggagtg | ttaggggggt | taaggacggg | 1440 |
| | gggggcggca | gctgggggct | actgcacgca | ccctttacta | aagccagttt | cctgggttctg | 1500 |
| 50 | atgggtattgg | ctcagttatg | ggagactaac | cataggggag | tggggatggg | ggaacccgga | 1560 |
| | ggctgtgcca | tctttgccat | gcccagtggt | cctgggcagg | ataatgtctc | agagatgccc | 1620 |
| | acgtctctgat | tcctccaaac | ctgtggacag | aacccgcccc | gccccagggc | ctttgcagg | 1680 |
| | gtgatctccg | tgaggaccct | gaggtctggg | atccttcggg | actacctgca | ggcccgaaaa | 1740 |
| | gtaatccagg | ggttctggga | agagggcggc | aggagggtca | gagggggggc | gcctcaggac | 1800 |
| 55 | gatggaggca | gtcagtctga | ggctgaaaag | ggagggaggg | cctcgagccc | agggctcaga | 1860 |
| | gcgcctccag | aagctggaaa | aagcggggaa | gggacctcc | acggagcctg | cagcaggaag | 1920 |
| | gcacggctgg | cccttagccc | accaggggcc | atcgtggacc | tccggcctcc | gtgccatagg | 1980 |
| | agggcactcg | cgctgccctt | ctagcatgaa | gtgtgtgggg | atttgcagaa | gcaacaggaa | 2040 |
| | acccatgcac | tgtgaatcta | ggattatttc | aaaacaaaag | tttacagaaa | catccaagga | 2100 |
| 60 | cagggctgaa | gtgcctccgg | gcaagggcag | ggcaggcag | agtattttta | tttagctatt | 2160 |
| | ttattttatt | tacttacttt | ctgagacaga | gttatgtctc | tggtgcccag | gctggagtgc | 2220 |
| | agcggcatga | tcttggtcca | ctgcaacctc | cgtctcctgg | gttcaagcaa | ttctcgtgcc | 2280 |
| | tcagctcccc | aagtagctgg | gatttcaggc | gtgcaccacc | acacccgggt | aattttgtat | 2340 |
| | tttttagtaga | gatgggcttt | caccatgttg | gtcaagctga | tctcaaaatc | ctgacctcag | 2400 |
| 65 | gtgatccgcc | cacctcagcc | ttccaaagtg | ctgggattac | aggcatgagc | cactgcacct | 2460 |
| | ggcctattta | accattttta | aacttccctg | ggctcaagtc | acacccactg | gtaaggagtt | 2520 |
| | catggagttc | aatttccctc | ttactcagga | tttaccctcc | tttgatattt | tctgtaattc | 2580 |
| | ttcgtagact | ggggatacac | cgtctcttga | catattcaca | gtttctgtga | ccacctgtta | 2640 |
| | ttccatggga | cccactgcag | gggcagctgg | gaggctgcag | gcttcagggt | ccagtggggt | 2700 |
| 70 | tgccatctgc | cagtagaana | ctgatgtaga | atcagggcgc | aagtgtggac | actgtcctga | 2760 |
| | atctcaatgt | ctcagtgtgt | gctgaaacat | gtagaatata | aagtccatcc | ctcctactct | 2820 |
| | actgggattg | agcccccttc | ctatcccccc | ccaggggcag | aggagttcct | ctcactcctg | 2880 |
| | tgagggaagg | aatgatactt | tggtattttt | cactgctggg | actgaatcca | ctggtttcatt | 2940 |
| | tggttggtttg | tttggttttg | tttgagaggg | ggtttctact | ttgttgctca | ggctggaggg | 3000 |
| | agtgcataatg | cgcgactctg | gcttactgca | gcctctgcct | cccagggttc | agtgaattct | 3060 |
| 75 | ctgcttccgc | ctcccatttg | gctgggatta | caggcaccgc | ccaccatgcc | cagctaatatt | 3120 |
| | tttgattttt | tagtagagac | gggggtgggt | gggggttcac | atgttggtcc | ggctgggtctc | 3180 |
| | gaacttctga | cctcagatga | ttccacctgc | tctgcctcct | aaagtgtctg | gattacaggt | 3240 |

007260-9428360

| | | | | | | |
|----|---|---|---|--|--|---|
| 5 | gtgagccacc gaagctcacc tgtagaaca acacactaac tgccgggagg ttccatttct aaccagtgt cctagtggca ggattttctag gagcgtgaca aatttctctc atttcagtgt tttctcgccc cagctgtcct tctactgctg gcttggaccc catctggcag ccggtgcgcg cccgggtggg gagaaactgc tgacgggagg tcgtcccccg ccggagcccg gcggcccaag cgggttaacc cctgcaccct ccgcgcggag gacgcccagg ccccttcacc gggtcccccg cgccctcttc ccctggcccc | atgcccagct ccactcaagt ctcttgatgt tgcacccata cgtttctctg tctcttccct agctacaact gagacaattc aagagcgacc gcccaggagg ggcagtttct ttgccgacct ccttagatcc gcggttctgc ggctgggaagt cgaggctgcc acagagtgc gccagcagga gattaacaga aaagagaaat cactccggga ccgcgtctac acgcccgcg ggctgcgcga cacagcctag gggagcgcca cagctgcgct accgcgctcc ttccagctcc cccagccccc tcgcgcgccg ggccaccccc | cagaattttac gttggtgtgt tttactactgt atactggggg ccatgcacat cttttaaaat taacttttgt acaaacacag tgtaatccta gggtgcgagg gaaagtagga cagctacaga aaacttgagg cggggcccca cgggcccctc ctccaccctg ggggcccagg gcgcctgggt tttgggtggg gacgggcctg ggctcccgct gcgcctccgt tccggacctg cgcacctgtt gccgatttga gcggcgcgcg gtcggggcca ccacgtggcg gcctcctccg tccggccctc agtttcaggc gcgatg | tctgttttaga tttaagccaa gatgactaag gtcttctggg gggtttaatt tggtttttct tggaacaaat ccctttaaaa agttatttaca ctgttcaaat aaggttacat atccctgcaa aaccgggagt ggcttgaggg agctctgcag tgccggcggg ttcaaggccg ccatttccca tttgcctcatg tgtcaaggag gcccgtccag cctccccctc gaggcagccc cccagggcct ccacatcatg ctggggccct cgccggccag ccagtggaat ggaccggggc gccccgtccc ccccttccct cctgctgcgc | aacatctggg tgatagaatt acatcatcag tatcagcaat actccagcat atgttggtct tttccaaacc aggcttaggg agacgaggct gctagctcca ttaagggtgc ggcctcggga ctggattcct ggaccagtgg tccgaggcct atgtgaccag ttgtggctgg cccaagtcgc ggagcaatgc acgtccggca tgggtctccg ccacatcatg ctggggccct cgccggccag ccagtggaat ggaccggggc gccccgtccc ccccttccct cctgctgcgc | tctgaggtag 3300 tttttattgt 3360 cttttcaaag 3420 cttcattgaa 3480 aatcttctgc 3540 ctctgcagag 3600 gcccccttgc 3660 atcataagg 3720 aacctccagg 3780 taaataaagc 3840 gtttgttagc 3900 gaccagaag 3960 gggaagtcc 4020 ccgtgtggct 4080 ggagcagggt 4140 atgttggcct 4200 tgtgaggcgc 4260 acgggaccgc 4320 ctcgcccgct 4380 ggggaagtgt 4440 gtcctcggtg 4500 ttcgtgtgct 4560 gatcaggcca 4620 gccccctcct 4680 cgctggcgct 4740 acccccgggt 4800 cgccgggcaca 4860 acccgtcctg 4920 gacccctccc 4980 tccgcggccc 5040 acgtgggaag 5100 5126 |
| 35 | <210> 2 <211> 4042 <212> DNA <213> Homo sapiens | | | | | |
| 40 | <400> 2 gtttcaggca cgatgccgag aggtgctgcc agcgcgggga gggagcgacg tggtggcccg gcttgcgctg gcagctacct tgctgcgcgg tgctgggtgg ctgcccactc aacgggacct gtgcgaggag gtggcgctgc gcaggagcgc aagaagccac gccgcagaca cttgtcccc agctgcggcc tcgtggagac cccgcctgcc acgcgcagtg ccccagcagc aggaggacac agggtgtacg ccaggacaaa atgccaaagt tgccgaggag tccttgccaa tcttttatgt tctggagcaa agctgtcgga gactccgctt tgggagccag cactgttccg tgctggccct aggacccgcc tccccagga gcgtgcgtcg | gcgctgcgct cgctccccgc gttggccacg cccggcggtc gcccggcccg agtgctgcag gctggacggg gcccacacag gcccagcgac cggtggcgac tcccagctgc ggcccggccc gaacctatgc gcgcgggggc ccctgagccg tggaaccagt ctcttttgag ccacgcgggc gggtgtacgc ctccttccca ctcagctctc catctttctg ccagcgctac cccctacggg cggtgtctgt agaccccgct cttcgtgcgg cgaacgcgcg ctcgtgcag cccagggtt gttcctgcac cacggagacc gttgcaaaag agcagaggtc catccccaa aacgtttccg cgtgtcaaac ggacgatata gcctgagctg caggctcacg gtatgccgtg | ctgctgcgca tgccgagccg ttcgtgcggc ttccgcgcgc cccgcccccc aggtctgtgc gcccgcgggg gtgaccgacg gacgtgctgg gcctaccagg cgccacacag gtcacccagg gtcacctgct tgctggggcc ctagtggacc ccgggggtccc gaagtctgcc ccgttgggca tctgtgtggt ctggcacgcg catcgcgggc acttctctca tgaggccccc cctggatgac ggccctctgt agacgcactg agccccaggg agctgctccg gcccggctgt acaccaagaa ggaagatgag cgcccgacga gtgtgtacgt agaacaggct gacagcactt gggaagccag tgccggccg ggggccgagc cgccggcgccc ggcgccacct aggttgatgt ccagcactat ccgcccattg | cctgggccccg gctgcgcagc ccagggtctg gtgcctgtgt gggtctctgc gaagaacgtg ggccttcacc gagcggggcg ggcacgctgc gccgctgtac ccgaaggcgt cctgggacct gttgcccagg ggggtccttg gtcacctgcc ccactcccac accacgtccc ctcctcaggc cctgactggc agggactccc tetggagctg cccgtgcga ctctgtggcg ccagcacagc gccccccagg gttcatctcc cgtgcgggac gcaccgtctg cgtcgagctg ctttttctac gaagagggtg gccccgcctg tgtgaacatg tctcacctcg cgccctcctg cgtgctgcgt gacgggcgcg caaaccaccg gcacgtcccg | gccacccccg 60 cactaccgcg 120 cggtgtgtgc 180 tgctgtccct 240 ctgaaggagc 300 ctggccttgc 360 accagcgctg 420 tgggggctgc 480 gcgctctttg 540 cagctcggcg 600 ctgggatgag 660 ccagccccgg 720 aggccccagg 780 gccccccggg 840 agaccccgcc 900 ccatccgtgg 960 tgggacacgc 1020 gacaaggagc 1080 gctcggaggc 1140 cgaggttgc 1200 cttgggaacc 1260 gctcgggtca 1320 gccccccagg 1380 agccccctgg 1440 ctctggggct 1500 ctggggaagc 1560 tgctgttggc 1620 cgtgaggaga 1680 ctcaggtctt 1740 cggaagagt 1800 cagctgcggg 1860 ctgacgtcca 1920 gactacgtcg 1980 agggtgaagg 2040 ggcgccctct 2100 gtgcggggcc 2160 tacgacacca 2220 aacacgtact 2280 aaggccttca 2340 | |
| 60 | | | | | | |
| 65 | | | | | | |
| 70 | | | | | | |
| 75 | | | | | | |

| | | | | | | | |
|----|---|--|---|--|---|--|---|
| 5 | agagccacgt tgaggagac aggccagcag tcaggggcaa tgctctgcag acgggtgtgt cgaaaacctt tgcggaagac ttcagatgcc tggaggtgca tcaaccgcgg tgaagtgtca acatctacaa catttcatca cctccctctg gcgcgcggg tcaagctgac agacgcagct acccggcact aggccgagag ggcggccac aggcctgcat aagggctgag caccgccagg tccatcccca accatccagg aggtgtgccc ttggggggag aaaaaaaaaa | ctctaccttg cagcccgcgt tgccctcttc gtcctacgtc cctgtgctac cctgcgtttg cctcaggacc agtgggtgaac ggcccacggc gagcgactac cttcaaggct cagcctgttt gatcctcctg gcaagtttgg ctactccatg ccctctgccc tcgacaccgt gagtcggaag gccctcagac cagacaccag acccaggccc gtccggctga tgtccagcac ccagcttttc gattcgccat tggagaccct tgtacacagg gtgctgtggg aaaaaaaaaa | acagacctcc agggatgccg gacgtcttcc cagtgccagg ggcgacatgg gtggatgatt ctgggtccgag ttccctgtag ctattcccct tccagctatg gggaggaaca ctggatttgc ctgcaggcgt aagaacccca ctgaaagcca tccgaggcgg gtcacctacg ctcccgggga ttcaagacca cagccctgtc gcaccgctgg aggctgagtg acctgcgctc ctcaccagga tgttccaccc gagaaggacc cgaggaccct agtaaaatac aa | agccgtacat tcgtcatcga tacgttccat ggatcccgca agaacaagct tcttgttggg gtgtccctga aagacgaggc gggtgcggct cccggacctc tgctcgcaaa ctgtgaacag acaggtttca catttttccct agaacgcagg tgcatgtggc tgccactcct cgacgtgac tcctggactg acgcccgggt gagtcgtagg tccggctgag ttcacttccc ggcccggcttc tcgcccctgcc ctgggagctc gcacctggat tgaatatatg | gcgacagttc gcagagctcc gtgccaccac gggtcccatc gtttgcccgg gacacctcac gtatggctgc cctgggtggc gctgctggat catcagagcc actctttggg cctccagacg cgcatgtgtg gcgcgtcact gatgtcgctg gtgccaccaa ggggtcactc tgccctggag atggccaccc ctacgtccca cctgagtgag gcctgagcga cacaggctgg cactccccac ctcctttggc tgggaatttg gggggtccct agtttttcag | gtgggtcacc tccctgaatg gccgtgcgca ctctccacgc attcggcggg ctcaccacag gtgggtgaact acggcctttt acccggaccc agtctcacct gtctttgcggc gtgtgcacca ctgcagctcc tctgacacgg ggggccaaag gcattcctgc aggacagccc gcccgagcca gcccacagcc gggagggagg tgtttggccc gtgtccagcc cgctcgggtc ataggaaatg ttccaccccc gagtgaccaa gtgggtcaaa ttttgaaaaa | 2400 2460 2520 2580 2640 2700 2760 2820 2880 2940 3000 3060 3120 3180 3240 3300 3360 3420 3480 3540 3600 3660 3720 3780 3840 3900 3960 4020 4042 |
| | <210> 3 <211> 11276 <212> DNA <213> Homo sapiens | | | | | | |
| 35 | <400> 3 acttgagccc tggtgacaga tcttctctgg atacaaacac ttaaaaaagg acccacggta tcaaaaaagt yccaaggcgg aaccttgtcg cagctactcg gagccgggat aaaaaaaaaa aaaagcaaga cagaaataaa ttttgaaaag acctaataaa caaaggatca aaaatagata agcccaaaca agagaagccc gaattccaat tctacatggc aacaacaaaa aaatcctcaa gtgatcaagt atgtgataca cagaaaaaag tatacaagaa aggccaaggt gagacctggg tagtcccagc tgcagtggag gaataagaag aggaggtgga tttcaacata agccttccct tagtactaga ctggaaagga gacttaagac caatgtacaa caaaaaagca tctacaatga | aagagttcaa atgagaccct ccacagtggg atgaaaatta aattgaaaaa tacagcaaaa agaaaagcca gcagatcgcc ctactaaaaa ggaggctgag tgcgccattg aagttagaaaa gcaaaactaaa tgaaaactgaa ataaacaaaa ataaagtcag ctagaggcta aattcctaga gaccaataac aggacccaat cctactcaaa cagtattacc cagaaagaaa caaaacacta gggattttatt tcattcccaac atttgataaa acatacaggc gggatgatgt ctacaaaaaa tagtctggag catgaacatg aaggagaagg ggagaagtgg ataaaaagccc ctaagatctg agtccctagc agaagtcaaa accactaaaa aaatcagtag gctacaaata aaactataaa | ggctacgggtg gtctcaaaaa acaaaaccag aacaatatac tttattttaag gcagtgctaa ggcgacgtgg tgaggctcagg tactaaaaat gcaggataac gactccagcc acttaaaaat cctaaaattg agataacaat ttgacaaaac agatgaaaaa ctatgagcaa tgcatacaa aataatggga ggcttccctg ctattctgaa ctgatttcaa gaaaactaca gcaaaccaaa ccagggatgg aaaatgaagt attctgcacc caggcacagt cttggggcca ctttttttaa gctgaggtgg tcactgtact agaagggaga aagggggaag tatatgacag gaaaatgaca agagcaatca ttatcctgtt aactatttag tattttctata aaattaaaca atgttgataa | agccatgatt aaaaaaaaaa aaatcaacaa ttctgaatga caaatgataa gaagggaagt ctcatgcctg agttcgagac gtggggcatg cgcttgaacc tgggttaacaa acaacctaat gtaaaagaaa acaaaagatc ttgtcccaga agagacatta ctgtacacta ctaccaagat ttaaagccat ctggatttta aaatagagga aaccagacaa ggccaatatc ttaaacaaca aaggatgggt acaaaacta cttcatgata ggctcacacc ggagtttgag aaattagcca gagaatcact ccagcctaga agggaggagg ggaagggaag accgaggtag agggccactg gataagagaa tgcatgatgt gctgaaattt ttccaacagc gctaggaatt aagaaattga | gcaacaccac aattgaaata caagaggaaat ccagtgaatg cggaaacata tatagctata taatcccagc cagcctgacc gtggcacatg caggaggtgg gagtgaaacc gatgcacctt agaaataata aacaaaatta ctaagaaaaa caactgatac ataaattgaa tgaacctatg aataaaaaag ccaatcattt aagaataact aaacacatca cctgatgaat ccttcgaaaag caacatatgc tatgattatt aaaaccctca tgcatcccca actagcctgg ggcatgatgg taagcctagg caacagaaca aagggaggag gaggagaag tattatgagg ttcaccactg agaaataaaa atgatcttat ggtacagcag aaacaatctg aaccacagaa agagggcaca | acgccagcct atataaagca tttgaaaact aatgaagaaa acctctcaaa agcagctaca actttgggag aacacagaga cctgtaacct aggttgcggt ctgtctcaag aaagaactag aagatcagag aaagtgtggt aggaaagagt cagagaaatt aaacctagaa agaaactcaa ctcctagcaa aaagaagaat ccaaactcat aaaacaaaaa actgatacaa atcattcatt aatcaatca tcactttatg aaaaccaggg gcactctggg gcaacaaaaa catatgcctg aggctcaggc agaccccact gaggagaagg aagaaacata aaaaactgaa gtattcaaat ggcatccaaa atctggaaaa gatacaaaa aaaaagaaac gtgaaagatc aaaaaagaaa | 60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 1680 1740 1800 1860 1920 1980 2040 2100 2160 2220 2280 2340 2400 2460 2520 |
| 75 | | | | | | | |

007260 9122360

| | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
| | agatattcca | tgttcataga | ttggaagaat | aaatactgtt | aaaatgtcca | tactacccaa | 2580 |
| | agcaattttac | aaatttcaatg | caatcccttat | taaaataacta | atgacgtttct | tcacagaaaa | 2640 |
| | agaagaaaca | atttctaagat | ttgtacagaa | ccacaaaaga | cccagaatag | ccaaagctat | 2700 |
| 5 | cctgaccaaa | aagaacaaaa | ctggaagcat | ccattacatt | gacttcaaat | tatactacaa | 2760 |
| | agctatagta | acccaaacta | catggtactg | gcataaaaac | agatgagaca | tggaccagag | 2820 |
| | gaacagaata | gagaatccag | aaacaaatcc | atgcatctac | agtgaactca | tttttgacaa | 2880 |
| | aggtgccaag | aacatacttt | ggggaaaaga | taatctcttc | aataaatggt | gctggaggaa | 2940 |
| | ctggatatcc | atatgcaaaa | taacaatact | agaactctgt | ctctcaccat | atacaaaagc | 3000 |
| 10 | aaatcaaaat | ggatgaaagg | cttaaatcta | aaacctcaaa | ctttgcaact | actaaaagaa | 3060 |
| | aacaccggag | aaactctcca | ggacattgga | gtgggcaaa | acttcttgag | taattccctg | 3120 |
| | caggcaacagg | caaccaaaagc | aaaaacagac | aaatgggac | atatcaagtt | aaaaagcttc | 3180 |
| | tgcccagcaa | aggaaacaat | caacaaagag | aagagacaac | ccacagaatg | ggagaatata | 3240 |
| | tttgcaaaat | attcatctaa | caaggaatta | ataaccagta | tatataagga | gctcaacta | 3300 |
| | ctctataaga | aaaacaccta | ataagctgat | tttcaaaaat | aagcaaaaaga | tctgggtaga | 3360 |
| 15 | cattttctcaa | aataagtcac | acaaatggca | aacaggcatc | tgaatgtgtg | ctcaacacca | 3420 |
| | ctgatcatca | gagaaatgca | aatcaaaact | actatgagag | atcatctcat | cccagttaaa | 3480 |
| | atggcctttta | ttcaaaagag | aggcaataac | aaatggcag | gaggatgtgg | ataaaagtaa | 3540 |
| | acccttggac | actgttgggtg | ggaatggaaa | ttgtaccac | tatggagaa | agtttgaaag | 3600 |
| | ttcctcaaaa | aactaaaaat | aaagctacca | tacagcaatc | ccattgtctag | gtatatactc | 3660 |
| 20 | aatcaggagg | aatcagtgta | tcaacaagct | atctccactc | ccacatttac | tgcagactgc | 3720 |
| | ttcatagcag | ccaaggtttg | gaagcaacct | cagtgtccat | caacagacga | atggaaaaag | 3780 |
| | aaaatgtggt | gcacatacac | aatggagtag | tacgcagcca | taaaaaagaa | tgagatcctg | 3840 |
| | tcagtgtgcaa | cagcatgggg | ggcactgggtc | agtatgttaa | gtgaaataag | ccaggcactg | 3900 |
| 25 | aaagacaaac | ttttcatgtt | ctcccttact | tgtgggagca | aaaattaaaa | caattgacat | 3960 |
| | agaaatatag | gagaatgggtg | gttctagagg | gggtgggggac | aggggtgacta | gagtcaacaa | 4020 |
| | taattttatg | tatgttttaa | aataactaaa | agagtataat | tgggttgttt | gttaacacaa | 4080 |
| | gaaaggataa | atgcttgaag | gtgacagata | ccccatttac | cctgatgtga | ttattacaca | 4140 |
| | ttgtatgcct | gtatcaaaat | atctcatgta | tgctatagat | ataaacccta | ctatatcaaa | 4200 |
| | aattaaaatt | ttaatggcca | ggcaggtggg | ctcatgtccg | taattcccag | actttgggag | 4260 |
| 30 | gccgaggcgg | gtggatcacc | tgaggtcagg | agtttgaaac | cagtctggcc | accatgatga | 4320 |
| | aacctgtgtc | ctactaaaga | tacaaaaatt | agccaggcgt | gggtggcacat | acctgtagtc | 4380 |
| | ccaactactc | aggaggctga | gacaggagaa | ttgcttgaac | ctgggaggcg | gaggttgcag | 4440 |
| | tgagccgaga | tcattgacct | gcactgcagc | ctgggtgaca | gagcaagact | ccatctcaaa | 4500 |
| 35 | acaaaaacaa | aaaaaagaag | attaaaaattg | taatttttat | gtaccgtata | aataataact | 4560 |
| | ctactatatt | agaagttaaa | aattaaaaca | attataaaag | gtaattaaac | gtattactta | 4620 |
| | aaataagaac | aatgtatgtg | gggtttctag | cttctgaaga | agtaaaagtt | atggccacga | 4680 |
| | tggcagaaat | gtgaggaggg | aaacagtgga | gttactgttg | ttagacgctc | atactctctg | 4740 |
| | taagtgaact | aattttaacc | aaagcacagg | tgaggagaagt | taaagaggca | ttctataagc | 4800 |
| 40 | cctaataacaa | ctgctaataa | tggtgaaaag | taactcttat | taattaccaa | taattacaga | 4860 |
| | tatctctaaa | atcgagctgc | agaattggca | cgctctgatca | caccgtctct | tcattcacgg | 4920 |
| | tgcttttttt | cttgtgtgtg | tggagatttt | cgtatgtgtg | ttcgtgtttg | gttaaaactta | 4980 |
| | atctgtatga | atcctgaaac | gaaaaatggg | ggtgatttcc | tccagaagaa | ttagagtacc | 5040 |
| | tggcaggaag | caggtgtgct | tgtgagacct | agccacttca | atcttcaagg | gtctctggcc | 5100 |
| 45 | aagaccaggg | tgcaaggcag | agggctgtatg | accagggac | aggaaagctc | ggatgggaag | 5160 |
| | gggcatgag | aagcctgcct | cggtgtgtgag | cagcgcatga | agtgccctta | tttacgcttt | 5220 |
| | gcaaaagattg | ctctggatag | catctggaaa | agggcgccag | cggaatgca | aggagtccga | 5280 |
| | agcctcctgc | tcaaacccag | gccagcagct | atggcgccca | ccggggcggtg | tgccagaggg | 5340 |
| | agaggagtca | aggcacctcg | aagtatggct | taaatctttt | tttccactga | agcagtgaac | 5400 |
| | aagggtgtatt | ctgagggaaag | cttgagttag | gtgccttctt | taaaacagaa | agtcattgaa | 5460 |
| 50 | gcacccttct | caagggaaaa | ccagacgcgc | ctctgcggg | catttactct | tttctctctt | 5520 |
| | ccctctcttg | ccctcgcggg | ttctgatcgg | gacagagtga | cccccggtga | gcttctccga | 5580 |
| | gcccgtgtcg | aggacctctt | tgcaaaaggc | ttccacagacc | cccgccctgg | agagaggagt | 5640 |
| | ctgagcctgg | cttaataaca | aactgggatg | tggctggggg | cggacagcga | cggcgaggat | 5700 |
| 55 | caaagactta | attccatgag | taaatccaac | ctttccacat | ccgaatggat | ttggatttta | 5760 |
| | tcttaatatc | ttcttaaaat | tcatacaata | acattcagga | ctgcagaaat | ccaaaggcgt | 5820 |
| | aaaacaggaa | ctgagctatg | tttgcacaag | ttcaaggact | taataaccat | gttcagaggg | 5880 |
| | atttttcgcc | ctaagtactt | tttattgggt | ttcataagg | ggcttagggg | gcaaggggaa | 5940 |
| | gtacacaggg | agaggcctgg | gcggcagggc | tatgagcag | gcaggggcac | cggggagaga | 6000 |
| 60 | gtccccggcc | tgggagctgc | acagcaggac | cactgaccgt | cctccctggg | agctgccaca | 6060 |
| | ttgggcaacg | cgaaggcggt | cacgctgcgt | gtgactcagg | acccataacc | ggcttccctg | 6120 |
| | gcccacccac | actaacccag | gaagtccagg | agctctgaac | ccgtggaac | gaacatgacc | 6180 |
| | cttgactgcc | tgcttccctg | ggtgggtcaa | gggtaatgaa | gtggtgtgca | ggaatggcc | 6240 |
| | atgtaaatga | cacgactctg | ctgattggga | ccgttctctc | catcattatt | catcttcacc | 6300 |
| 65 | cccaaggact | gaatgattcc | agcaactctt | tcgggtgtga | caagccatga | caaaactcag | 6360 |
| | tacaaaacac | actcttttat | taggcccaca | gagcacggsc | cacacccctg | atatattaa | 6420 |
| | agtccaggag | agatgaggct | gctttcagcc | accaggctgg | gggtgacaaca | gcggtgtaac | 6480 |
| | agtctgttcc | tctagactag | tagaccctgg | caggcactcc | cccagattct | agggcctggg | 6540 |
| | tgctgtcttc | cgagggcgcc | atctgcccct | gagactcagc | ctgggggtgc | acactgaggc | 6600 |
| 70 | cagccctgtc | ttccacacct | ccgctccag | gcctcagctt | ctccagcagc | ttcctaaacc | 6660 |
| | ctgggtgggc | cgtgttccag | cgctactgtc | tcacctgtcc | cactgtgtct | tgtctcagcg | 6720 |
| | acgtagctcg | cacggttctc | cctcacatgc | ggtgtctgtc | tccttcccca | acactcacat | 6780 |
| | cggttgaaag | gaggagattc | tgcyctctcc | agactggctc | ctctgagcct | gaacctggct | 6840 |
| | cgtggccccc | gatgcagggt | cctggcgctc | ggctgcacgc | tgacctccat | ttccaggcgc | 6900 |
| 75 | ttcccgtctc | ctgtcatctg | ccggggcctg | ccgggtgtgt | cttctgtttt | tgtgtctctt | 6960 |
| | ttccagctca | gctgcgtgtg | tctctgccc | ctagggtctc | gggggttttt | taggcatagg | 7020 |
| | acggggggcgt | gggtgggccc | ggcgctcttg | ggaaatgcaa | catttgggtg | tgaagtagg | 7080 |
| | agtgctctgc | ctcacctagg | ttcacgggca | caggcctggg | gatggagccc | ccggcaggga | 7140 |
| | cccgcctctc | tctgcccagc | actttctctg | ccccctccct | ctggaacaca | gagtgggcag | 7200 |

SECRET

| | | | | | | | |
|----|--|---|--|---|---|--|--|
| 5 | tccacaagc gccccacagc ccgacccccg tttaacaaac gaggaacatg tgccacctcc aggggagtgg ccttttacta cataggggag ctctgggcagg aaccgcgccg atccttctggg aggagggtca ggaggggaggg gggacctctc atcgtggacc gtgtgtgggg aaaaaagg ggcaggcacg ggtatgctct cgtctctctg gtgcaccacc gtcaagctga ctgggattac ggctcaagtc gttacctctc catattcaca gagggctgca atcagggcgc gtagaataat ccaggggcag cactgtgggt ggtttcactc gcctctgcct caggccaccg ggggttccct ctgtcttctc tctgtttaga tttaagccaa gtgactaaag gtccttctgg ggtgttaatt tgtttttctt tggaacaaat ccctttaaaa agtatttaca ctgttcaaat aaaggttatc atccctgcaa aaccgcgagt ggtctggagg agctctgcag tgcggggcgg gtcaaggcgc ccattttccc tttgctcatg tgtcaaggag gcccgctcag cctccccctc gaggcagccc cccaggcctc cctctctccg ggcggggaag ggccgggctc gagggactgg cgcggacccc cccagccctc agcgcgtcgt | actaagcatc cctgggaatt ctgttttatt tggttaaacg ccgtttataa atgggatacg ttaggggggt aagccagttt tggggatggg ataatgtctt gccccagggc actacctgca gaggggggca cctcgagccc acgggagcgt tcggccctcc at ttgcagaa tttacagaaa agtgatttct tgttgcccag gttcaagcaa acaccgcgt tctcaaaatc aggcatagc acaccactgc tttgatattt gtttctgtga gcttcaggtc aagtgtggac aagtcctatc aggagtctct actgaatcca ttgttgtctc cccaggttca ccaccatgcc atgttgcca aaagtcctcc aacatctggg tgatagaatt acatcatcag tatcagcaat actccagcat atgttgactt tttccaaacc aggcttaggg agacgaggtc gctagctcca ttaaggttgc ggcctcgcca ctggattctc ggaccagtgg tccgagcgtt atgtgaccag ttgtgtgtcg tctgttctcg gtggggcccc cccaagtcgc acgtctcgga tgggtctccg ccatgcatg ctggggccct cctctctccg ggcggggaag ggccgggctc gagggactgg cgcggacccc cccagccctc agcgcgtcgt | ctcttcccaa cacgtgacta ttaatagcta taacgggtcca agcctgcagg tacgcaacat taaggacggt cctggttctg ggaacccgga agagatgcct ctttgcaggt ggcccgaaaa gcctcaggac aggcctgcaa cagcaggga gtgccatagg gcaacaggaa catccaaggaa ttatagctatt gctggagtgc ttctcgtgcc aat11t1gtat ctgacctcag cactgcacct gtaaggaggt tctgtaattc ccactgttta ccagtggggg actgtcctga ctctactctc ctcactcctg ctgtttcatt ggctggaggg agtgattctc cagctaattt ggctgggtct gattacaggt tctgaggtag tttttattgt cttttcaaag cttcattgaa aatcttctgc ctctgcagag gcccctttgc atcactaagg gtttgttagc gaccagaag gggaagtctc ccgtgtgggt ggagccaggt atgttggcct tgttagggcg tgtgaggcgc acgggaccgc ctcgcgcctc ggggaagtgt gtcctcgggt ttcgtgtgtc gatcaggcca gcccctccct cgctggcgct accccgggt cgcgggcaca accggtcctg gaccctctcc tcggcgccc acgtgggaag | aagaccagc cgacatcat caaaagcagg tccgcacggt catctcaagg gtcaaaaaa gggggcggca atgggtattg ggctgtgcca agctcctgat gtgatctccg gtaatccagg gatggaggca gcgcctccag gcaaggctgg aggggcactgc accatgcac cagggctgaa ttattttatt agcggcatga tcagcctccc tttagtaga gtgatccgcc ggcctattta gtggagttc ttcgtagact tcccattgga tcctatctgc atctcaatgt actgggattg tgagggaagg tgttgggttg agtgcaatgg ctgcttccgc tttgtatttt gaacttctga gtgagccacc gaagctcacc tgttagaaca acacactaac tgccgggagg ttccatttct aaccagttga cctagtggca ggattttctag gagcgtgaca aatttctctc at ttcaggtg ttctctgcc cagctgtcct tctactgtgc gcctggacc catctgccag ccggtgcgcg ccgggtgggt gagaacctgc tgcagggagg tcgtccccag ccggagcccg gcggccaaag cgggttacc ctgtgacctc cgcccgagg gacgcccag cccttccacc gggtcccccg cgccctctcc cctggcccc | attggcaccc gtacacactc gaacacctgc ggacagttcc gaattacgct aaagaatttc gctgggggct ctcagttatg tctttgccat tcccccaaac tgaggacctc ggttctggga gtcagttctga aagctggaaa cccttagccc cgctgcccct tgtgaattct gtgctcccg tactactctt tcttggtcca aagtagctgg gatgggcttt cacctcagcc accattttaa aatttcccct ggggatacac ccactgcag cagtagaaaac ctcagttgtg agccctctcc atcgtacttt tttgttttgt cgcgactctg ctcccattg tagtagagac cctcagatga atgcccagta ccactcaagt ctctgtatgt tgcaccatta cgtttctctg tctcttccct agctacaact gagacaattc aagagcgacc gccagggag ggcagtttct ttggcgacct ctctagacct acagagtgc gccagcagga gattaaacaga aaagagaaat cactccggga ccgcgtctac acgcccccg ggctgcgcga cacagcctag gggagcgaga cagctgcgct accgcgtctc ttccagctc cccagcccc tcggcgcgcg ggccaccccc | ctggacattt ccgtccacga tcaaatgtcc tcacagttaa gagtcaaaac accccatggc gtgcacgca ggagactaac gcccaggtgt ctgtggacag gaggtctggg agagggcggg ggctgagaaa aagcggggaa accagggccc ctagcatgaa ggattatttc gcaagggcag ctgtgacaga ctgcaacctc gatttcaggc caccatgttg tcccaaagtg aacttcccct ttactcagga cgtctcttga gggcagctgg ctgaatgtag gctgaaacat ctatcccccc tgttattttt tttgagaggc gcttactgca tctgggatta gggggtgggt tccacctgcc cagaatttac gttgtgggtg ttacactgt atactggggt ccatgcacat ctttttaaatt taacttttgt acaaacacag tgttaactta ggtgcgaggc gaaagttaga cagctacagc aaacttgagc cggggcccca cgggcctcct ctccacctg ggggcccagg gctcctggct tttggggtgg gacgggcctg ggtcccgcgt gcgctcctg tcgggacctg cgacctgtt gccgatttga gcgggcgcg gtcggggcca ccacgtggcg ctcctctccg tcggggccct agtttctaggc gcgatg | 7260 7320 7380 7440 7500 7560 7620 7680 7740 7800 7860 7920 7980 8040 8100 8160 8220 8280 8340 8400 8460 8520 8580 8640 8700 8760 8820 8880 8940 9000 9060 9120 9180 9240 9300 9360 9420 9480 9540 9600 9660 9720 9780 9840 9900 9960 10020 10080 10140 10200 10260 10320 10380 10440 10500 10560 10620 10680 10740 10800 10860 10920 10980 11040 11100 11160 11220 11276 |
| 70 | <210> 4 <211> 104 <212> DNA <213> Homo sapiens | | | | | | |
| 75 | <400> 4 gtgggctctc catgcggaga | ccgggggtcgg cgacgcgagg | cgctccggctg cgactcaggg | ggggttagggg cgcttcccc | cgcccggggg gcag | | |
| | | | | | 60 104 | | |

```
<210> 5
<211> 8616
<212> DNA
<213> Homo sapiens
```

5

[illegible][illegible]

DECEMBER

| | | | | | | | |
|----|-------------|--------------|-------------|-------------|-------------|-------------|------|
| | ggctggagtg | taattggcaca | atctcggttc | actgcaacct | ctgcctcttc | ggttcaagca | 4380 |
| | gttctcatte | ctcaacctca | tgagttagctg | ggattacagg | cgcccaccac | cacgcctggc | 4440 |
| | taatttttgt | aatttttagta | gagataggett | ttccacctgt | tggccaggct | ggctctaaac | 4500 |
| 5 | tectgacctc | aatgtacatcg | ccgccttggt | ctctccacag | tgtctgggatt | acaggtgcaa | 4560 |
| | gccaccgtgc | cgggcatacc | ttgatctttt | aaaaatgaagt | ctgaaacatt | gctacccttg | 4620 |
| | tectgagcaa | taagaccttt | agtgatattt | agctctggcc | acccccagc | ctgtgtgctg | 4680 |
| | ttttcccctg | tgacttagtt | ctatctcagg | catcttgaca | cccccaaac | ctaagcataa | 4740 |
| | ttaatatgtt | tttccgtgtt | gagtggttct | gtagctttgc | ccccgcctg | cttttctctc | 4800 |
| | tttgttcccc | ctgtgtcttc | tgctctcagg | ccgcgctctg | gggtccctct | ccttgtctct | 4860 |
| 10 | tgcgtgggtt | ttctgtcttg | ttattgtctg | taaaccacag | ctttacctgt | ctgtggcttc | 4920 |
| | atggcatcta | gcgacgtccg | gggacctctg | cttatgatgc | acagatgaag | atgtggagac | 4980 |
| | tcacaggagg | ggcgggtcac | ttggcccggt | atgtgtctgga | gcaccacgtg | gccagcggtc | 5040 |
| | cttagccagt | gagtgacagc | aacgtccgct | cgcctctgggt | tcagcctgga | aaacccagg | 5100 |
| 15 | catgtcgggg | tctggtggct | ccgcggtgtc | gagtttgaia | tcgcgcaaac | ctgcggtgtg | 5160 |
| | gcgcacagtc | tacgggtgct | gcctgcgggg | ggagtgctctg | ctctctccct | tctgcttggc | 5220 |
| | aaccaggaca | aaggatgagg | ctccgagccg | ttgtcgccca | acaggagcat | gcagtgagcc | 5280 |
| | atgtggataa | ttttaaaatt | tctaggctgg | gcgcggtggc | tcacgcctgt | aatcccagca | 5340 |
| | ctttggggagt | ccaagggcgg | tggatcacga | ggtcaggagg | tcgagaccat | cttggccaac | 5400 |
| 20 | atgatgaaac | cccatctgta | ctaaaaaacac | aaaaattagc | tggcggtgggt | ggcggtgtcc | 5460 |
| | tgtaatccca | gctactcggg | aggctgaggc | aggagaattg | cttgaacctg | ggagttggaa | 5520 |
| | gttcgcagtga | gcgcacattg | caccactgca | ctccagcctg | gcaacacagc | gagactctgt | 5580 |
| | ctcaaaaaaa | aaaaaaaaaa | aaattctagta | aaattctagta | gccacattaa | aaagataaaa | 5640 |
| | aagaaaaggt | gaaattaatg | taataataga | ttttactgaa | gccacagcat | tccacacctc | 5700 |
| 25 | atcatttttag | gggtgttatgt | gtggggagcat | cactcacagg | acatttgaca | ttttttgagc | 5760 |
| | tttgtctcgc | ggatccccgt | tgtagctccc | gtgcgtggcc | atctcggcct | ggacctgtctg | 5820 |
| | ggcttcccat | ggccatggct | gttgtaccag | atggtgcagg | tccgggatga | ggtcgccagg | 5880 |
| | ccctcagtg | gctggatgtg | cagtgctcgg | atggtgcagg | tctgggatga | ggtcgccagg | 5940 |
| | ccctgctcgt | agctggatgt | gtggtgtctg | gatggtgcag | gtcaggggtg | aggtctccag | 6000 |
| 30 | gccctcgggt | agctggaggt | atggagtccg | gatgatgcag | gtccgggggt | aggtcgccag | 6060 |
| | gccctgctgt | gagctggatg | tgtggtgtct | ggatggtgca | ggtcaggggt | gaggtctcca | 6120 |
| | ggccctcggg | aagctggagg | tatggatccc | ggatgatgca | ggtcaggggt | gaggtctcca | 6180 |
| | ggccctgctg | tgagctggat | gtgtggtgtc | tggatggtgc | aggtctgggg | tgaggtcacc | 6240 |
| | aggccctcgc | gtgagctggg | tgtgcggtgt | cttgatgggtg | caggtctgga | gtgaggtcgc | 6300 |
| 35 | cagacgggtg | cagaccatgc | ggtagctggg | atatcgcggt | tccggatggt | gcaggtctg | 6360 |
| | ggtgaggttg | ccaggccctg | ctgtgagttg | gatgtggggg | gtccggatgc | tgcaggtccg | 6420 |
| | gtgtgaggtc | accaggccct | ctgtgagctg | ggatgtgtgg | tgtctggatg | gtgcaggtct | 6480 |
| | ggggtgaagg | tcgcaggccc | cctgcttgtg | agctggatgt | gtggtgtctg | gatggtgcag | 6540 |
| 40 | gtctggagtg | aggtcgccag | gccctcgggt | gagctggatg | gtcgggtgtc | ggatggtgca | 6600 |
| | gtccgggggtg | aggtcgccag | ggccctcggg | gagctggatg | tatggagtcg | ggatgggtgc | 6660 |
| | ggctcggagt | gaggtcgcca | gaccctcgtg | tgagctggat | gtcgggtgtc | tggatggtac | 6720 |
| | ggtcgggggt | tgaggtcgcc | agaccctgtg | gtgagctgga | tatcggtgtg | ccggatgggt | 6780 |
| | aggctcggag | gtgaggtctc | caggccctcg | gtgagctgga | ggtagggagt | ccgagatgat | 6840 |
| 45 | caggtcaggg | gtgaggtctc | caggccctgc | tgtgaactgg | atgtgcccgg | tctggatggg | 6900 |
| | gtaggtcggg | gggtgtggctg | ccaggccctc | ggtagactgg | aggtatggag | tccggatgat | 6960 |
| | gcaggtcggg | gggtgaggtc | ccaggccctg | ctgtgagctg | gatgtgcggc | gtctggatgg | 7020 |
| | tgcaggtctg | gggtgtggct | gccaggccct | cggtagactg | gaggtatgga | gtccggatga | 7080 |
| | tgaggtctcg | gggtgaggtt | gccaggccct | gctgtgagct | ggatgtgtctg | tatccggatg | 7140 |
| 50 | gtgcagtcg | gggtgaggtc | caccaggccc | tgcggtgagc | tggttgtgctg | gtgtccgggt | 7200 |
| | gtgcaggtct | ggggtgaggt | tcgccaggcc | ctcggtagac | tggatgtgctg | gtgtcccggt | 7260 |
| | gtcgcaggtg | cggggtgagt | gggtgaggtc | gctaggccct | tggttggctg | gatgtgcccgt | 7320 |
| | gtccggatgg | tgcaggtctg | gggtgaggtc | gccaggccct | tgggtgagctg | gatgtgcccgt | 7380 |
| | gtctgcagtg | tgcaggtctg | gggtgaggtc | gccaggccct | tgggtgagctg | gatgtgcccgt | 7440 |
| 55 | gtccggatgg | tgcaggtctg | gggtgaggtc | gccaggccct | tgggtgagctg | gatgtgcccgt | 7500 |
| | gtctgcagtg | tgcaggtctg | gggtgaggtc | gccaggccct | tgggtgagctg | gatgtgcccgt | 7560 |
| | gtccggatgg | tgcaggtctg | gggtgaggtc | gccaggccct | tgggtgagctg | gatgtgcccgt | 7620 |
| | tgtctggatg | gtcaggtctc | ggggtgaggt | agccaaagcc | ttcggtagac | tgatgtgagg | 7680 |
| | gtgtccggat | gggtgcaggtc | cggggtgagg | tcgcaggccc | ctcgggttag | ctggatgtgc | 7740 |
| | gggtgtccgga | ttggtgcaggt | ccggggtgag | gtcaccaggc | cctgcgggtta | gctggatgtg | 7800 |
| 60 | cgggtctctgg | atggtgcagg | tcgggggtga | gggtcgcagg | ccctgtctgtg | agctggatgt | 7860 |
| | gctgtatccg | gatggtgcag | gtccggggtg | aggtcgccag | gcctctcagt | gagctggatg | 7920 |
| | tgtctgtatcc | ggatgggtgca | gggtctggcgt | gaggtcgcca | ggccctgcgg | ttagctggat | 7980 |
| | atgcgggtgtc | ggatgggtgca | gggtctggcgt | gaggtcacca | ggccctgcgg | ttagctggat | 8040 |
| | gtgcgggtgtc | cggatgggtgc | aggtctgggg | tgaggtcgcc | aggccctgtc | gtgagctgga | 8100 |
| 65 | tgtgtctgtat | ccggatgggtg | caggtctcggg | gtgaggtcgc | caggccctgc | ggtgagctgg | 8160 |
| | atgtctctgta | tccggatgggt | gcaggtctgg | ctgtgaggtg | ccaggccctg | cgggtgagctg | 8220 |
| | gatgtgcagt | gtacgggatgg | tgcaggtccg | gggtgaggtc | gccaggccct | gcggtggggt | 8280 |
| | gtatgtgtgt | tgtcttggatg | gtgcaggtcc | ggggtgagtt | tcgcaggccc | tgcggttgagc | 8340 |
| | tgtgatgtgtg | gtgtctggat | gtgcaggtcc | cgggtgtgagt | cggccaggcc | ctcggtagagc | 8400 |
| | tggatatgcg | gtgtccccgt | gtccgaatgg | tgaggttcca | gggtgaggtc | gccaggccct | 8460 |
| 70 | tgttggggctg | gatgtgcgggt | gtccggatgg | tgaggtctg | gggtgaggtc | gccaggccct | 8520 |
| | gggtgagctg | gatgtgcgggt | gtccggatgg | tgaggtctg | gggtgaggtc | accaggccct | 8580 |
| | cggtgatctg | gatgtggcat | gtccttctctg | tttaag | | | 8616 |
| 75 | <210> | 6 | | | | | |
| | <211> | 2089 | | | | | |
| | <212> | DNA | | | | | |
| | <213> | Homo sapiens | | | | | |

<400> 6
 gtactgtatc cccacgccag gcctctgctt ctggaagtc tggaaacacca gcccgccctc 60
 agcatgcgcc tgtctccact tgcctgtgct tccctggctg tgcagctctg ggctgggagc 120
 cagggggccc gtcacagggc tgggtccaagt ggattctgtg caaggctctg actgacctga 180
 gctcacgttc tcttacttgt aaaaatcagga gtttctgcca agtgggtctct aggggtttgt 240
 aagcagaagg gatttaaatc agatggaaac actaccacta gcctccttgc ctttccctgg 300
 gatgtgggtc tgattctctc tctctttttt ttttcttttt tgagatggag tctcactctg 360
 ttgcccaggg tggagtgcag tggcataact ttggctcaact gcaacctcca cctcctgggt 420
 ttaagcgatt caccagcctc agcctcctaa gtactgtgga ttacaggcac ctgccaccac 480
 gcttggctaa tttttgtact tttaggagag acggggtttc accatgttgg ccaggctggg 540
 ctggaactca tgacctcagg tgatccaccc accttggcct cccaaagtgc tgggtttaca 600
 ggctaagcca ccgtgcccag ccccgatctc tcttttaatt catgctgttc tgtatgaatc 660
 ttaaatctat ttgattttagg tcatgagagg ataaaaatc accacttgg ccactcactg 720
 cagggagcac ctgtgcaggg agcacctggg gataggagag ttccaccatg agctaacttc 780
 taggtggctg catttgaatg gctgtgagat tttgtctgca atgttcggct gatgagagtg 840
 tgaagtgtgt acagattcaa gctggatttg catcagtgag ggacgggagc gctggctctg 900
 gagatgccag cctggctgag cccaggccat ggtattagct tctccgtgtc ccgcccaggc 960
 tgactgtgga gggcttttagt cagaagatca gggcttcccc agtccccctg cacactcgag 1020
 tccctggggg gccttgggac accccatgcc ccaaatcagg atgtctgcag agggagctgg 1080
 cagcagacct cgtcagagg taaacacagct ctgggctggg gaccccgacg tgggtgctgg 1140
 gccattttct tgcactctgg ggagggtcag ggctttccct gtgggaacaa gtaatacac 1200
 aatgcacctt acctagactt tacacgtatt taatgggtgt cgacccaaca tggctcattg 1260
 accagtattt tggaaagaat ttaattgggg tgaccggaag gaggagacag acgtgggtgg 1320
 ccccaagatg ctctcttgtc ctactgggag cctgtgtctg cctggggggc ctgggagggc 1380
 cctcctccct ccgcagggtg ggacgggttc ccgtctcttt tctactctgc tggctgggtc 1440
 agggcaccag ctccggagca cccgcgccc cagtgtccac ggagtgccag gctgtcagcc 1500
 acagatgccc aggtccagg tgggcccgtc gactccccct gcccccattg ctgtgttttg 1560
 gggaaaaggc caagggcaga ggtgtcagga ggtgtcaggg cactatgagag ctgattctgc 1620
 tcttggctg agctgccctg agcagcctct cccgcctct ccacttgaag ggatgtggct 1680
 cttttacctt gggggtcttg cctggggcca gccttgggtt accccagtgg cctgtaccaga 1740
 gggacaggca tctctgtgtg aggggcatgg gttcacgtgg cccagatagc agctctggag 1800
 caggctccct ggtgctgatg gtgggacagt caccctgggg gttgaccgcc ggactggggc 1860
 tcccagggtt tgactatagg accaggtgtc cagggtgccct gcaagtagag gggctctcag 1920
 aggcgtctgg ctggcatggg tggacgtggc cccggcatg gcttcacgc tgtgctgccg 1980
 tgggtgccct gagccctcac tgagtctgtg ggggtctgtg gcttcccgtg agcttcccc 2040
 tagtctgttg tctggctgag caagcctcct gaggggctct ctattgcag 2089

<210> 7
 211> 687
 "<212> DNA
 213> Homo sapiens

<400> 7
 gtggctgtgc tttgggttaa ctctctttt aaacagaagt gcgtttgagc cccacatttg 60
 gtatcagctt agatgaagg cccggaggag gggccacggg acacagccag ggccatggca 120
 cggcgccaac ccatttgtgc gcacagttag gtggccgagg tgcgggtgcc tccagaaaag 180
 cagcgtgggg gtgtaggggg agctcctggg gcagggacag gctctgagga ccacaagaag 240
 cagccggggc agggcctgga tgcagcacgg cccgaggtcc tggatccgtg tctgtctgtg 300
 gtgcgcagcc tccgtgcgct tccgcttacg ggcgccggg accaggccac gactgcaggg 360
 agcccaccgg gctctgagga tctctggacct tgccccacg ctctgcacc ccacccctgt 420
 ggtgcggggt gctgcgggtg ccccgctatc tgaggagagt gtgggggtg gtggacagag 480
 gtttggcatg aggtaccctg gtgcaacaca catgcggcca ggaacctgt tcaaacagg 540
 tctgaggaag ctgggagggg ttctaggtcc cgggtctggg tggctgggga cactggggag 600
 gggctgcttc tcccctgggt cctatgggtg ggggtgggac ttggccggat ccactttcct 660
 gactgtctcc catgtgtgac ccgacag

<210> 8
 <211> 494
 <212> DNA
 <213> Homo sapiens

<400> 8
 gtgggtgcgc gggaccccgc tgagcagccc tgctggacct tgggagtggt tgcttgattg 60
 gcacctcatg ttgggtggag gaggtactcc tgggtggggc gcaggggagt cagggtgacc 120
 tgtcactggt gaggacacac ctggcaccta ggggtggagg cttcagcctt tctgacagca 180
 catggggccc actgtgcacc ctgactgccc gggctctcat tcccaaggag ggtcccact 240
 gattccagtt tccgtgcagc aaggaaccgc aacgctctag ccaccaggcc ccggtgctct 300
 gcacccaggt cctgagccag ggggtctcct tctttagggt cagagagggg acacagccc 360
 ccttgccttc ggggtctgga gtgggtgggg tcagagagag agtgggggg accgcaggc 420
 caggccctga gggcagaggt gatgtctgag ttctgcgtg gccactgtca gtctcctgc 480
 ctccactcac acag

<210> 9
 <211> 865
 <212> DNA
 <213> Homo sapiens

<400> 9

| | | | | | | | |
|----|-------------|-------------|-------------|-------------|------------|------------|-----|
| 5 | gtaagggtcca | cgtgtgatag | tctgtgtccag | gatgtgtgtc | tctgggatat | gaatgtgtct | 60 |
| | agaatgcagt | cgtgtctctgt | atgcgtttct | gtgggtggagg | tacttccatg | atttacacat | 120 |
| | ctgtgatatg | cgtgtgtggc | acgtgtgtgt | cgtgggtgcat | gtatctgttg | cgtgcatatt | 180 |
| | tgtgtgtgtg | gtgtgtgtgg | cacgtgtgtg | tccatggtgt | gtgtgcctgt | ggtgtgcatg | 240 |
| | tgtgtgtgtc | tgtgtacacgt | gcatgttcat | gctgtgtgct | gcatgtctgt | gatgtgccta | 300 |
| | tttgtgtgtg | gtgtgtgcat | gtgtccgtga | catatgcgtg | tctatggcat | gggtgtgtgt | 360 |
| | ggcccttgg | ccttactcct | tcttctcca | ggcatggtcc | gcaccattgt | cctcacgtc | 420 |
| 10 | tgggtgtctg | gtttggggag | ctccacattc | agggctctca | cttctagcat | gggtgcccc | 480 |
| | gtcctgtcac | agggtctggc | cttggagact | gtaagccagg | tttgagagga | gagtagggat | 540 |
| | gctgggtgta | ccttctctga | cccttggcac | ccccaggacc | ccagtctggc | ctatgccggc | 600 |
| | tccatgagat | ataggaaggc | tgattcaggc | ctcgtctccc | gggacacact | cctcccagag | 660 |
| | cgcccggggg | ccttggggct | cggcagggtg | gaaaggggcc | ctgggcttgg | gttcccaccc | 720 |
| | agtgtgcatg | agcacgctgg | aggggttaagc | cctcaaagtc | gtgccaggcc | ggggtgcaga | 780 |
| 15 | ggtgaagaag | tatccctgga | gcttcggctc | ggggagaggc | acatgtggaa | acccacaagg | 840 |
| | acctctttct | ctgactttct | gagct | | | | 865 |

<210> 10

<211> 3782

<212> DNA

<213> Homo sapiens

<400> 10

| | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|------|
| 25 | tgtgggattg | gttttcatgt | gtgggatagg | tggggatctg | tgggattggg | ttttatgagt | 60 |
| | ggggtaaacac | agagtttcaag | gagagctttc | ttcctgtagt | gggtctgcag | gtgtcccaac | 120 |
| | agcttttattg | aggagaccat | atcttctctt | gaactatggg | cgggtttata | gtaagtacag | 180 |
| | ggtgtggagg | cctccctctg | gtctcctgtt | ctgtttcttc | cactctgggg | tcgtgtgggt | 240 |
| | cctgtctgtg | tgtgtggccg | gtgggcaggg | cttccaggcc | tccttgtgtt | catggcctg | 300 |
| 30 | gatgtggccc | tggctaggct | cgtctctgtg | aattccctcg | cgagtggag | gctttcttct | 360 |
| | ttttcttttt | tctttctttt | tttttttttt | tgataacaga | gtctcgctct | tttttgccta | 420 |
| | ggctggagtg | gtttggcggt | atcttggctc | actgcaacct | gtgtctctcg | agttcaagca | 480 |
| | attctcttgc | ctcagcctct | caagtagctg | gaattatagg | cgcccaccac | catgctgact | 540 |
| | aattttttgta | atttttagtag | agacgaggtt | tctccatggt | ggccaggctg | gtctcgaaact | 600 |
| 35 | cctgacctca | ggtgatctct | ccacctcggc | ctcccaaggt | gctgggatga | cagggtgtgaa | 660 |
| | ccgcgcgcgc | cggcgcagac | tcgcttctct | cagcttccgt | gagatctgca | gcgatagctg | 720 |
| | cctgcagcct | tgggtgctgac | aacctccgtt | ttccttctcc | aggtctcgct | aggggtcttt | 780 |
| | ccatttcatg | actctcttca | cagaagagtt | tcacgtgtgc | tgatttcccg | gctgtttctc | 840 |
| | cggttaattgg | gtctctgtgt | ttatcgatgg | ctctcctcca | tttcttttag | gctttgttta | 900 |
| 40 | ttgttgtttt | tccggctcct | tgaaggaaaa | gtttcgatta | tggatgtttg | aactttcttt | 960 |
| | tctaacaacg | catctgaagt | tgccgttttc | cctctaagac | agggatcccc | agggccctgg | 1020 |
| | ctgtggagtg | gcaccgctct | ggggcctgtt | ggaaaccggg | cgcacagcgg | gaggctaggt | 1080 |
| | gggggtgtggg | gagccagcgt | tcccgccctga | gccccgcccc | tctcagatca | gcagtgccat | 1140 |
| | gcggtgctca | gagggcgaca | caccctactg | agaactgtgc | gtgagagggg | tctagattct | 1200 |
| 45 | gtgtctctta | tgggaattca | atgctctgat | atctgagggt | gaaccgtttg | ctccaaaaac | 1260 |
| | catccctctc | cccactgtct | tctgttgtaa | aaatcgctct | ccacgaaacc | agtccctggg | 1320 |
| | accacaatgg | ttggggaccc | tgtgtctaaag | acctgtctca | gcagcctctc | gtcagtggtg | 1380 |
| | atatattggc | tttctgtgtg | tgagtccaga | ataattacgg | atttctgtga | tgctttcccg | 1440 |
| | cgacctcaga | cccattgggt | atttgtgggc | gtgttgccct | ctcctgggtt | gggaagggtg | 1500 |
| 50 | caggccccat | gtaccttctc | gttactgcct | tccagggttg | ttctcagggt | tgaatcgta | 1560 |
| | tcgatgtggg | tttagccac | ggcctctgat | ccagctcctg | ggggctgggg | aacatgctga | 1620 |
| | agcacagagt | caccgtgcgc | gtcttttgat | gcctcacaag | ctcgaggcct | cctgtgtccg | 1680 |
| | tgttagtggt | tgtcacgtgc | ctgtctcact | cctgtcttgg | ggacgcaggg | gcttagcagg | 1740 |
| | tcccgtagta | aatgacaagc | gtcctggagg | agctctcaga | ataggagggt | gggggtcccg | 1800 |
| 55 | tctctctccc | gcgtcttccg | actcttctcc | tgctgtgtgt | gtggctgcac | ctgcacccct | 1860 |
| | gcaatccctc | cagcactggg | ctggagaggc | ccgggagctc | gagtgccact | tgtgcccagt | 1920 |
| | gactgtggat | ggcagtcggt | cacgggggtc | tgatgtgtgg | tgactgtgga | tggcgggttg | 1980 |
| | tcacaggggt | ctgatgtgtg | gtgactgtgg | atggcggtcg | tggggtctga | tgtgggtgta | 2040 |
| | gtggatggcg | gtcgtggggg | ctgatgtgtg | gtgactgtgg | atggcggtcg | tgggggtctga | 2100 |
| 60 | tgtggtgact | gtggatggcg | gtcgtggggg | ctgatgtggg | gactgtggat | ggcggctcgt | 2160 |
| | gggtctgatg | tgggtgactgt | ggatggcagt | cgtgggggtc | gatgtgtggg | gactgtggat | 2220 |
| | ggcgggtcgt | gggtctgatg | tgggtgactgt | ggatggcagt | cgtgggggtc | gatgtgtggg | 2280 |
| | gactgtggat | ggcgggtcgt | gggtctgatg | tgtgtgact | gtggatggcg | gtcgtggggg | 2340 |
| | ctgatgtgtg | gtgactgtgg | atggcggtcg | tggggtctga | tgtgtgggtga | ctgtggatgg | 2400 |
| 65 | cggtcgtggg | gtctgatgtg | gtgactgtgg | atggcggtcg | tggggtctga | tgtgtgggtga | 2460 |
| | ctgtggatgg | tgatcggtea | caggggtctg | atgtgtgggt | actgtggatg | gcggctcgtg | 2520 |
| | ggtctgatgt | gtgggtgactg | tggatgggtg | tcggtcacag | gggtctgatg | tgtgggtgta | 2580 |
| | tgtggatggcg | gtcgtggggg | ctgatgtgtg | gtgactgtgg | atggcggttg | gtcccggggg | 2640 |
| | ctgatgtgtg | ggtgactgtg | gatggcgatc | ggtcacaggg | gtctgatgtg | tgggtgactg | 2700 |
| 70 | ggatggcggt | cgtgggggtc | gatgtgtggg | gactgtggat | ggcgggtcgt | gggtctgatg | 2760 |
| | tgtgggtgact | gtggatggcg | gtcgtggggg | ctgatgtggg | gactgtggat | ggcgggtcgt | 2820 |
| | gggtctgatg | tgggtgactgt | ggatggcggt | cgtgggggtc | gatgtgtggg | gactgtggat | 2880 |
| | ggcgggttgg | cccgggggtc | tgatgtgtgg | tgactgtgga | tggcggtcgt | gggggtctgat | 2940 |
| | gtgggtgactg | tggatggcag | tcgtgggggtc | tgatgtgtgg | tgactgtgga | tggcggtcgt | 3000 |
| 75 | gggtctgatg | gtgtgggtgac | tgtggatggc | ggctgtgggg | tctgatgtgt | ggtgactgtg | 3060 |
| | gatggcggtc | gtgggggtcgt | atgtgtggatg | actgtggatg | gcgggtcgtg | ggtctgatgt | 3120 |
| | ggtgactgtg | gatggcggtc | gtgggggtcgt | atgtgtgggtg | actgtggatg | gtgatcgggtc | 3180 |
| | acaggggtct | gatgtgtggg | gactgtggatg | ggcgggtcgt | gggtctgatg | tgtgtgtgact | 3240 |
| | gtggatggcg | gtcgtggggg | ctgatgtggg | gactgtggat | ggcgggtcgt | gggtctgatg | 3300 |

| | | | | | | | |
|----|--|--|--|---|--|--|---|
| 5 | tgtggtgact gtcacagggg actgtggatg tgatgtgtgg tcgggtcacag actttgcgtc tgggtctcat agtgtccagc ag | gtggatggcg tctgatgtgt gcgggtcggtg tgactgtgga gggtctgatg ctcggccccc cccggccatc tctggccggg | gtcgtagggt gggtgactgt gggtctgatg tggcgggtcgt tgtgttagct cggcccccgt ggcttggccg gcaggccaca | ctgatgtgtg gatggcggtc gtgggtgactg gggggtctgat gcagggtggag tccccaaaca cagggtccaca tttgtggctc | gtgactgtgg gtgggggtctg tggtatggcgg gtggtgactg tcccagggtg gaagcttccc cgtcctgatc atgccccttc | atggcagtcg atgtgtgggtg tcgtgggggtc tggtatgggtg gtctgtagct agggcgtctc ggaagaaaca ctctgcccgg 3782 | 3360 3420 3480 3540 3600 3660 3720 3780 |
| 10 | <210> 11 <211> 980 <212> DNA <213> Homo sapiens | | | | | | |
| 15 | <400> 11 | | | | | | |
| 20 | gtctggggcag aatcactggg tgtgatgggg tgcgacagct ccgagtgccc caaaggaaag agctggcccc cgggcccgtg ggccccgtgg cctgttagca ctctggcgga ccagaccctg cgggtgggctg gtgtcacaca gatggccctg cggagggtct tctcccgtct | tgccccgcag ctcatgaccg gcagtatgag gctgcattca tttgttcatg gtgtccccct tcagtgtgtg ttgagccacg ccctttgcag cttgcctggc atttccctgg tgccccgcag tgtgggtgtg ctctgcctaa cattccagcc tggccacgtg gcttttcgca | gggtggggcag gacagactgt ctgtgtgcct ggcacctgct atlttgctaaa ccttttaggag gtctgaggcc ccccgctgag atgtgggtctg tctaggggac ctcccagggt ctgggagca agccccagctg gccccatgtg cagccccgca gtccctgctg | ggactcccag tggccccggg tggcgaaatc cacgtttgac tgtcttctct ggcaggccat aaaggaaacg cgggcccctc tccacgtggc agtcgtgtcc gggggtggag actcctggat gacccacagg gtctgcagag cttcatcaca tctcagcacc | cagtgggtcc gggcagtggtg tgagctgggc tgccgggccc gccagttttg gtttgagccg tgtccccctt agtgtgggtg cctgtggctc accgcatgag gtggcctggg cacatatgcc tggtcccagag actcggcccc aacactgacc caccggctca | tccccctgggc gggaatgagc catgccaggc ctctccagtt atcttgaggc tgtcctgccc cttaggagga ctgtcccagc tttgagatg gctcagagac ctgtggggac atccgggcca gagacgttct gtccagccac ccaaaaggga ctccccatgtg | 60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 980 |
| 35 | <210> 12 <211> 2485 <212> DNA <213> Homo sapiens | | | | | | |
| 40 | <400> 12 | | | | | | |
| 45 | gtgagtcagg tgctcacctc ttttctggcc ggctcggtct gggggtgtgga tgccgagcagc cggctctcac tctctgccct cccatctgga ccatggggga cgggtgttaga cacacctccc ggaggaaatt ggaggcctct tttatttaaa attataatat cacaaattgc taagcggccc cgggcctcct gtggcagggc ggtgactgtg tggtccagtt acagagagag ggggtggcgg gccccatcact ttttttgaga actcactgca gctgagatta gggtttttgc ctcggcctcc ctttttaagg gctgtcaagc tcgcgtggca actgtttgtc gtcactctga catgagctct tcattccggg aaagaaaacc | tgccccagggtg tctcctgccc cccggccccct gcggcagccg gttgcctcctg gtttgagcct acgcttgtat gtcgtgtgac aagtgcgggg ggcggccctg ggcaccagtg ggcaggcatc cgtgcacatc ctctgggagt aatataaacta ttattaaagt acatggcagc ccaggccccc tcgtgggtcgt tttgggggaa tctgtcctgt tggtccctgt tttccccatc gactcctaga gtgatatctg cggaaacgtca acctccgccc caggcaccca catgtttggc caaagtgtctg tgaccaccta tctcttagca gccatgcctt tgaaaacgca aactaggggc tcaccgtgga tcaagtgtct ttgatgattc | ccattgccct cttccccact ccggctcctg gagcggagca cgtggaggac gcagcttgtc ctctctctcc ccccgcgagg ttgaccgtgt gagagctgcc ctgggtgccac tgccctgcgac caaggctcatc gtctccagcg ttaattattg ataattagaa agagtgaatt agaatttcgt gaattttatt gtgaggtgat ccctaggaca ataaaaaactg catgtgtctca gttgggtgctg caccagcaag ctgtttgtctg ccccgggttc ccccctgcgc aggctgggtc ggattacagg tagcgcttcc acaggagtgg ctgtgtgcac cccttggcat aaggttgtat caaattcctt gggtctgtga agagcaaggga | gcgggtgggt gnccttctgc ggctgcaggc ggtgccacac gaggggcggg agctccaagt cgatacaaaa gcgcgggtct agtttgctcc gtcacacagc atcacgtcct cctgtgtgtg agcaaggctca gataaaaggac cattataagt atattaagta ttggccgagg gacaaaagtc aagatggatc gactgcgtcc cgacaggccc cttcaaaaacc caggggcgta gtgcttctgt gcaaaaagtg gttcttttct agtgaggctg agcatttctc ctggctaatt cgaactcctg tgtgagccat cgaaaaataac cgctcctgtg cttttaggttc ccttgttttg ccgttggcgc gaaaaaaaaa ataaactcta tgtgtgcaca | ggggcgggctg ccggggccac tcccagggcc gagggcctgga gggtgtgtct tactactgac ggattttatc ttctctctgt tctcgggggg cactgggtga ctggatttta cctggggaga tccgcagtc tgtgcacagc aatcactaat gtacacacgt gacacgtgtg cctccccaga aagtcacgta tcatgcccctg cgaagctcta tggtgcccc tctgtttgcg gcaaaaagtg tttcttttct agtgaggctg ctgcctcagc tttgtatttt acctcagggtg cacgcccagc aggtccttgt gctctgggga cacgggggcta agagtctctg gcagcgggcta aaaggagctc agatttaaga cctgtggctg | gcagggccttc cagagctctcc ccggaaacat aatggcaagc gggtcagggtg gctggacacc cgattctcat gactagattt cctgtgggtg gccacactca agtaaaacca gtggtagcac ggtggaacgt ttcgggaagc ggtatcagca tctggaaaaa cacatgtgtg gaagccacca ccgtccacgt acagacagga gtccccatcg aaaactaaga ttgactcgct cagtcctctt ttctttttct cgcgatctca ctcccagaca tagtagagag atccacccac cggaaagcct tttgagtagg tggctgaggg ttctgtctct cttctcgttg catgtagggt gggttaagcat aaccttaagt gatctgtttc | 60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 1680 1740 1800 1860 1920 1980 2040 2100 2160 2220 2280 |

001260" 5422360

| | | | | | | | |
|----|--------------------|-------------|-------------|-------------|------------|-------------|------|
| | agccgcccca | gtgcatgggtg | agagtgggga | gcagggtattg | tttgttcaga | ggtctcatct | 2340 |
| | ggtatgtttc | tgaggtgttt | gccggctgaa | tggtagacgt | gtcgtttgtg | tgatagaggt | 2400 |
| | tctgtgtctg | tgtgtggctc | ggtttgagtg | tacgcatgtc | cagcacatgc | cctgcccgtc | 2460 |
| 5 | tctccactgt | gtcttccgcg | cccg | | | | 2485 |
| | <210> 13 | | | | | | |
| | <211> 1984 | | | | | | |
| | <212> DNA | | | | | | |
| 10 | <213> Homo sapiens | | | | | | |
| | <400> 13 | | | | | | |
| | gtgaggcctc | ctcttcccca | ggggggcttg | ggtgggggtt | gatttgcttt | tgatgcatte | 60 |
| | agtgttaata | ttcctgtgtc | tctggagacc | atgactgtct | tgctctgagg | aaccagacaa | 120 |
| 15 | ggttgacagc | ccttctctgt | atgaagccgc | acggggaggg | ttgcacagcc | tgaggcagtc | 180 |
| | gggctccacc | caggtctctg | ccagcggcca | tgctccagag | cctcagggtg | cagcaggcgc | 240 |
| | gagggccgct | gcccctgcat | atgagcatgt | gaattcaaca | ccgaggaagc | acaccagctt | 300 |
| | ctgtcaccgt | acccaggttc | cgttagggtc | cttggggaga | tggggctggt | gcagcctgag | 360 |
| | gccccacatc | tcccagcagg | ccctcgacag | gtggcctgga | ctgggcgcct | cttcagccca | 420 |
| 20 | ttggccatcc | cacttgcatg | gggtctacac | ccaaggacgc | acacacctaa | atatcgtgcc | 480 |
| | aacctaatagt | ggttcaactc | agctggcttt | tattgacagc | agttaccttt | ttttttttaa | 540 |
| | tactttaagt | tctagggtac | atgtgcacga | cgtgcaggtt | agttacatat | gtatacatgt | 600 |
| | gccatggttg | tgtgtctcac | ccattaactc | atcattttaca | ttaggtatat | ctcctaatag | 660 |
| | tatccctccc | cactccccc | atcccatgac | aggccctggt | gtgtgatgtt | ccccaccctg | 720 |
| 25 | tgtccaagtg | ttctcattgt | tcagttccca | cctgtgagtg | agaacatgtg | gtgtttggtt | 780 |
| | tgctttccct | gcaatagtgt | gctcagagtg | atggttttcca | gcttcgtcca | tgttccctaca | 840 |
| | aaggacatga | actcatcctt | ttttatgact | gcatagttat | ccgtgtgtga | tatgtgccac | 900 |
| | atttttcttaa | tccagttctat | catacatgga | catltgggtt | ggttgcaagt | ctttgttact | 960 |
| 30 | gtgaatagtg | ccgaataaaa | catacgtgtg | catgtgtctt | tatagacgca | tgatttataa | 1020 |
| | tcctttgggt | atataccag | taatgggatg | gctgggtcaa | atgggtattc | tagttctaga | 1080 |
| | tccttgaggga | atcacccac | tgtcttccac | aatggttgaa | ctagtttaca | ctcccaccaa | 1140 |
| | cagtgtaaaa | gtgtttcgtt | ctgggagagg | atgtggacag | cagttatttt | tttatgaaaa | 1200 |
| | tagtatcact | gaacaagcag | acagttagtg | aaggatgcgt | caggaagcct | gcaggccaca | 1260 |
| | cagccatttc | tctcgaagac | tccgggtttt | tcttgtcat | cttttgaac | cttagctcca | 1320 |
| 35 | attatagcat | tgtacagtga | tccaggtttc | tcttacctaa | ggttcaggtt | ctagattgaa | 1380 |
| | ataagtttat | gtaacagaaa | caaaaatttc | ttgtacacac | aacttgcctc | gggatttgga | 1440 |
| | ggaagagtgc | ctcgagctgg | cggcacactg | gtcagccctc | tgggacagga | tacctctggc | 1500 |
| | ccatggtcat | ggggcgctgg | gtctgggcct | gaggggtcac | cagtgacaca | tgcccagctt | 1560 |
| | cctgtggata | ggatctgggt | ctcggatcat | gctgaggacc | acagctgcca | tgtctggtaa | 1620 |
| 40 | gggcaccacg | tggctcagag | ggggcgagggt | tccagacccc | agctttctta | ccgtcttcaa | 1680 |
| | ttatttttcc | ctaagagtct | gagaagtggt | gccgcgcctg | atggccttcg | ttcgtcttca | 1740 |
| | gctggcacag | aattgcacaa | gctgtggtga | aacactgagt | acttataatg | aatgaggaat | 1800 |
| | tgctgtagca | gttaactgtg | gagagctcgt | ctgttggaag | gaattttaag | tttttctatt | 1860 |
| 45 | aaccgctttg | gagaatgtta | ctttattttt | ggctgtgtaa | attgtttgac | attcagttcc | 1920 |
| | tcgtagacag | atactacgta | aaaagtgtaa | agttaacctt | gctgtgtatt | ttcccttatt | 1980 |
| | ttag | | | | | | 1984 |
| | <210> 14 | | | | | | |
| | <211> 1871 | | | | | | |
| | <212> DNA | | | | | | |
| 50 | <213> Homo sapiens | | | | | | |
| | <400> 14 | | | | | | |
| | gtgaggcccg | tgcggtgtgt | ctgtggggac | ctccacagcc | tgtgggcttt | gcagttgagc | 60 |
| 55 | cccccggtgc | ctgccccgtg | caccgcagcg | ttgtctctgc | caagtcctct | ctctctgccg | 120 |
| | gtgctggatc | cgcaagagca | gaggcgcttg | gccgtgcacc | caggcctggg | ggcgcagggg | 180 |
| | caccttcggg | agggagtggg | taccgtgcag | gcccttggtc | tgcagagacg | caccagaggt | 240 |
| | acacacgtgg | tgagtgcaag | cggtagacctg | gctcctgtcg | ctctttggaa | agtcaagagt | 300 |
| | ggcggtctct | ggggccccag | tgagaccccc | aggagctgtg | cacagggcct | gcaggggcga | 360 |
| 60 | gctcggcagc | tcttccccag | gggtgcacctg | agcctgcgga | gagcaggagc | tgctgagtga | 420 |
| | gtggcccac | agcgttgcct | cggtgcacgt | tcttgcgttg | ggttgtttgg | gatcgggtgg | 480 |
| | agaatttgga | tttgtctagt | gctgtctgtc | tgaaccacgg | agatggctag | | |



<210> 16
 <211> 880
 <212> DNA
 <213> Homo sapiens

5

<400> 16
 gtgagcaggc tgatgggtcag cacagagtcc agagtccagg aggtgtgtgc gcaagtatgt 60
 gtgtgtgtgt gtgcgcgcgt gcctgcaagg ctgatgggtga ctggctgcac gtaagagtgc 120
 acatgtacgc atatacacgt gaggacatac atgtgtgcat gtgtgtacat gaaggcatgg 180
 cagtgtgtgc acaggtgtgc aagggcacaa gtgtgtgcac atgcgaatgc acacctgaca 240
 tgcatgtgtg ttcgtgcaca gtcgtgtggg cattcacgtg aggtgcatgc gtgtgggtgt 300
 gcagtgtgag tagcatgtgt gcacataaca tgtattgagg ggtcctcgtg ttcaccccg 360
 taggtcctca gcaccagtgc cactccttac aggatgagac ggggtcccg gccttggtgg 420
 gctgaggctc tgaagctgca gccctgaggg cattgtccca tctgggcatc cgcgtccact 480
 cctctccttg tgggcttctg tgtccactcc cctctcctg tgggcattta catccactcc 540
 actcctctct tctctgtggg atccgcgtcc actccctctc tctgtgggca tctgctgcca 600
 cctccctctc ctgtgggcat ttgctgctcc tccctctctc ggttctctcc tgtcttggcc 660
 gagcctcggg ggcaggcaga tgacacagag tcttgactcg cccaggggtg ttcgcagctg 720
 ccgggtgagg gccaggccgg atttacttgg gaagagggat agtttcttgt caaaatgttc 780
 ctcttctctg ttccatctga atggatgata aagcaaaaag taaaaactta aaatcccaga 840
 gaggtttcta ccgttttcta ctcttttctg gcgactctag 880

<210> 17
 <211> 3186
 <212> DNA
 <213> Homo sapiens

25

<400> 17
 gtgagccgcc accaaggggt gcaggccag cctccaggga cctccgcgc tctgtccacc 60
 tctgaccggg ggcttcacct tggaaactct ggggttttagg ggcaaggaat gtcttacgtt 120
 ttcagtgggt ctgctgcctg tgcacagtcc tgttcgcgtg gctctgtgca aagcacctgt 180
 tctccatctc tgggtagttg taggagccgg tgtggcccca ggtgtcccca ctgtgctgt 240
 gcaactggcg tgggacgtca tggaggccat cccagggcag caggggcatg gggtaaagag 300
 atgtttatgg ggagtccttg cagaggaggg tgggaaggtg tctgaacagt agatgggaga 360
 tcagatggcc ggaggatttg ggggtcagc aaagagggcc gaggtgggtg caggtagagg 420
 tgcgtggccc caccctcggg aaggtgcagc agagctgtgg ctccccacac agcccgcca 480
 gcacctgtgc tctgggcatg gctgtgtctc tggaaacttc cctgtcctgg ctggtcaggg 540
 ggtgccccct ccaagaatcg acaactttat cacagagggg agggccaatc tgtggaggcc 600
 acagggccag cttctgcctg gagtcagggc aggtgggtgg acaagcctcg gggctgtacc 660
 aaagggcagt cgggcaccac agggccgggc ctccacctca acaggcctcc cgagccactg 720
 ggagtcgaat gccaggagc cgaagcctc gcccatgag ggctgagaag gagtgtgag 780
 atttgtgtta cccaggcccg aggtgcgcg aattaccgtg cacacttgat gtgaaatgag 840
 gtcttcgtct atcgtggaaa cccagcaagg gctcacggga gatttttcca ttacaaggtc 900
 gtaccatgaa aatggttttt aaccgagtg cttgcccctt catgctctgg caggaggggc 960
 agagccacag ctgcatgtta ccgcctttgc accagctcca gaggtctggg accaggctgt 1020
 ctacgttcca ggggtgcgtc ggctcagacc gccctcctct ctgccttctc tctctgctc 1080
 aaatcttccc tctgttgcag cctcctgggg cgtgcctggg ccctcgtgca agctgctga 1140
 ctcttttccc gaaaccttgg ggggtgtgct gatacaggtg cactgagga ctggagggtg 1200
 ctgacactgt ggttgacccc aggttccagg tggcgtgctt ggggctcctc tgggccatga 1260
 tgaggtcaga ggagtttttc caggtaaaaa cctctgggaa actcccaggg ccattgtagc 1320
 tgccacctgc tcttcccata ttcagctcag tcttgcctc atttcccac cagggtctct 1380
 agctccgagg agctcccgtg gaggccctgg gctcagggca gggcggtga gtttcccac 1440
 ccattgtggg acccttgggt agtcgcttga ttgggtagcc ctgaggaggc cgagatgca 1500
 tgggcccagg gccgtttcca aacacagagt caggcacgtg gaaggcccag gaatcccctt 1560
 ccctcgaggc aggagtgagg gaacggagag ctgggccccg atttccaggc agccaggctg 1620
 cagtggggcg ggcgtgtgtg tccacgtgg cgtggggggc ggggtctgat tcaaatccgc 1680
 tggggctcgg ccttctctgg ccgtgctggc cgcgcctcca cacgggcttg ggggtggacg 1740
 cccgacctct agcagggtggc tatttctccc tttggaagag agccctcac ccattgattt 1800
 tgttccctc ctgggtcagg agcgtggccg tgtggcaacc cggggacett aggttattt 1860
 atttgtttta aaacattctg ggcctggcct ccgttgttgc taaatgggga aaagacatcc 1920
 cacctcagca gaggttactg gaggctgaaa ccgggtgtgt ggcttgactg gtgtgatctc 1980
 aggtcattcc agaagtggct caggaagtca gtgagaccag gtacatgggg ggctcaggca 2040
 gtgggtgaga tgaggtagac ggggggctca ggcagtgggt gaggccaggc acatgggggg 2100
 ctacggcact gggtagatg aggtacacgg ggggctcagg cagagggtca gaccaggtag 2160
 acgggggctc tgatcacacg ccatatgtgag catgtgtgt ttcatgttag 2220
 ccagggtctg gcacacctgc cccaaagtcc caggaaagct agaggccaaa gatggaggct 2280
 gacagggctg gcgcgggtggc tcacacctgt agtcccagca ctttgggagg ccgaggcgag 2340
 aggtccctc gagccaggga gtttaagacc agcctgagca acatagtaga accctactc 2400
 tatgaaaaat aaaaaaaaaa attagctgaa catggtgtgt tgcgcctgta gttccaatac 2460
 ttggggaggct gaagtgggag gatcacttga gccaggaggg tggaaagctg agtgagctga 2520
 gattgcacca caaatgcagt tcttgggaaa gaaacattta gtaggaactt aacctacaca 2640
 agaagactga cagaagccaa gtcggtgtct cgggtgcagt gcgatgagat gatgggtcct cacaccatca 2700
 cccagagccc agggtttatg caccacaggg cagggtggct cagaagggat gcgagagag 2760
 ttgatatacg atgacatcaa ggttgtctga cgaagggcag gattcatgat aagtacctgc 2820
 ttgtacacaa ggaacaatgg ataaactgga aaccttagag gccttcccgg aacaggggtc 2880
 aatcagaagc cagcatgggg ggtctgcac caggatggag ctgcttcagc ctccacatgc 2940
 gtgttcatac agatgggtgca cagaaacgca gtgtacctgt gcacacacag acacgcagct 3000

007260-34233560

| | | | | | | | |
|----|--------------------|-------------|-------------|--------------|-------------|-------------|------|
| | actcgcacac | acaagcacac | acacagacat | gcattgcattgc | atccgtgtgt | gtgcacctgt | 3060 |
| | gcccatgagg | aaacccatgc | atgtgcattc | atgcacgcac | acaggcaccg | gtgggcccac | 3120 |
| | gccacacccc | acgagcaccg | tctgattagg | aggcctttcc | tctgacgctg | tccgccatcc | 3180 |
| 5 | tctcag | | | | | | 3186 |
| | <210> 18 | | | | | | |
| | <211> 781 | | | | | | |
| | <212> DNA | | | | | | |
| 10 | <213> Homo sapiens | | | | | | |
| | <400> 18 | | | | | | |
| | gtatgtgcag | gtgcctggcc | tcaatggcag | cagtgcctgc | ctgctgggtgt | tagtgtgtca | 60 |
| | ggagactgag | tgaatctggg | cttaggaagt | tcttaccctt | tttcgcatca | ggaagtgggt | 120 |
| 15 | taacccaacc | actgtcaggc | tcgtctgccc | gccctctcgt | ggggtgagca | gagcacctga | 180 |
| | tggaggggac | aggagctgtc | tgggagctgc | catccttccc | accttgctct | gcctggggaa | 240 |
| | gcctctgggg | gcctgggtctc | tctgttttgc | cccatgggtg | gatttggggg | gcctggcctc | 300 |
| | tctgttttgc | cctgtgtgtg | gattgggctg | tctcccgctc | atggcactta | gggcccctgt | 360 |
| | gcaaacccag | gccaagggct | taggaggagg | ccaggcccag | gctacccccc | ccctctcagg | 420 |
| 20 | agcagaggcc | gcgtatcacc | acgacagagc | cccgcgctgt | cctctgcttc | ccagtcaccg | 480 |
| | tcctctgccc | ctggacactt | tgtccagcat | caggagggtt | tctgatccgt | ctgaaattca | 540 |
| | agccatgtcg | aacctgcggt | cctgagctta | acagcttcta | ctttctgttc | tttctgtgtt | 600 |
| | gtggaaattt | cacctggaga | agccgaagaa | aacattttctg | tcgtgactcc | tgcgggtgctt | 660 |
| 25 | gggtcgggac | agccagagat | ggagccaccc | cgcagaccgt | cgggtgtggg | cagctttccg | 720 |
| | gtgtctcctg | ggagggggagc | tgggctgggc | ctgtgactcc | tcagcctctg | ttttccccc | 780 |
| | g | | | | | | 781 |
| | <210> 19 | | | | | | |
| | <211> 536 | | | | | | |
| | <212> DNA | | | | | | |
| 30 | <213> Homo sapiens | | | | | | |
| | <400> 19 | | | | | | |
| 35 | gcaagtgtgg | gtggaggcca | gtgcggggccc | cacctgcccc | ggggtcatcc | ttgaacgccc | 60 |
| | tgtgtggggc | gagcagcctc | agatgctgct | gaagtgcaga | cgcccccggt | cctgacctgt | 120 |
| | ggggcctgga | gccacgctgg | cagccctatg | tgattaaacg | ctggtgtccc | caggccacgg | 180 |
| | agcctggcag | gggtccccaac | ttcttgaaac | cctgcttccc | atctcagggg | cgatgtctcc | 240 |
| | ccacgcttgg | gagccttctg | acccttgacc | tgtgtctctt | cacagcctct | tccttggtctg | 300 |
| 40 | ctgccctgag | ctcctggggg | cctgagcaag | ttcttctccc | gccccggcgc | tccagcgtca | 360 |
| | ctgggctgcc | tgtctgtctg | ccccggtgga | ggggtgtctg | tcctttcact | gaggttccca | 420 |
| | ccagccaggg | ccacgaggtg | caggccctgc | ctgcccgggc | acccacacgt | cctaggaggg | 480 |
| | ttggaggatg | ccacctcttg | cctcttcttg | aacggagtct | gattttggcc | ccgcag | 536 |
| | <210> 20 | | | | | | |
| 45 | <211> 3179 | | | | | | |
| | <212> DNA | | | | | | |
| | <213> Homo sapiens | | | | | | |
| | <400> 20 | | | | | | |
| 50 | atctcatgtt | tgaatcctaa | tgtgactgct | atagacacca | ctgtatgcaa | ttacagaagc | 60 |
| | ctgtgagtga | acgggggtgt | ggtcagtgcg | ggcccatggc | ctggctgtgc | atttacggaa | 120 |
| | gtctatgagt | gaatggggtt | gtggctcagt | cgggcccatt | gcctggctgg | gcctggggag | 180 |
| | ttctctgagt | tgtgaggcag | gaggggaagg | agggtagggg | atagacagtg | ggagccccca | 240 |
| 55 | ccctggaaga | cataacagta | agtccaggcc | cgaagggcag | cagggatgct | gggggcccag | 300 |
| | cttggggcgg | ggggatgatg | gagggccttg | ccagggtggc | agggatgatg | ggggccccag | 360 |
| | ctgggggtgg | aggggtgatg | gggggggctg | gtctgggtgg | cggggaagat | ggggaagcct | 420 |
| | ggctggggccc | cctcctcccc | tgcctcccac | ctgcagccgt | ggatccggat | gtgcttccct | 480 |
| | gggtcacatc | ctctggggca | tcagctttca | tggaggtggg | gggcaggggc | atgacaccat | 540 |
| 60 | cctgtataaa | atccaggatt | cctcctcctg | aacgccccaa | ctcaggttga | aagtacatt | 600 |
| | ccgcctcttg | ccatttctct | aagagttagc | caggattctg | atctctgaag | gggtggtagg | 660 |
| | gtggggcagt | ggaggggtgt | gacacaggag | gcttcagggt | ggggctgggt | atgctctctc | 720 |
| | atcctcttat | catctcccag | tctcatctct | catctcttta | tcattctccc | gtctcatctg | 780 |
| | tcttctctct | atctcccagt | ctcatctgtc | atctctttac | catctcccag | tctcatctct | 840 |
| 65 | tatctcttta | tctcttagtc | tcattccagc | ttactctcca | gggctgggtg | caggctcgca | 900 |
| | gtggagctgg | acatacgtcc | ttcctcaggg | agaaggaaat | ggaaggattg | cagagaacac | 960 |
| | gagggggcgg | tcagagggac | gcagtcttgg | gggtgaagaa | cagccccctc | tcagaagttg | 1020 |
| | gcttggggcca | cacgaacccg | agggccctgc | gtgagtggct | ccagagcctt | ccagcaggtc | 1080 |
| | cctgggtggg | ccttatggta | tggccgggtc | ctactgagtg | caccttggac | aggctctctg | 1140 |
| 70 | gtttgagtgc | agcccggacg | tgcctgggtg | cggggtgggg | gcttatggcc | actggatatg | 1200 |
| | gcgtcattta | ttgctgtctg | ttcagagaat | gtctgagtga | ccgagcctaa | tgtgtatggt | 1260 |
| | gggcccgaagt | ccacagactg | tgtcgtaaat | gcactctggt | gcctggagcc | cccgatagg | 1320 |
| | agctgtgagg | aaggaggggc | tcttggcagc | cggcctgggg | gcgcctttgc | cctgcaaaact | 1380 |
| | ggaaggggagc | ggccccgggc | gcccgtgggg | gacgacctca | agtgagaggt | tggacagaa | 1440 |
| 75 | agggccgggga | cttcccaggc | gcagagggcg | gtctcaggc | acacctgggt | ttgaatcac | 1500 |
| | gacccaacagg | tcaggccatt | gttcagctat | ccattcttcta | caaagctcca | gattcctgtt | 1560 |
| | tctccgggtg | ttttttgttg | aaattttact | caggatttact | tatatTTTTT | gctaaagtat | 1620 |
| | tagacccttta | aaaaaggtat | ttgctttgat | atggcttaac | tcactaaaga | cctactttat | 1680 |
| | ttgtctgttt | ttattttatta | ttattattat | tattagagat | ggtgtctact | ctgtcaccca | 1740 |

007260-24225560

| | | | | | | | |
|----|-------------|------------|------------|-------------|------------|------------|------|
| | ggttgtagt | gcagtggcac | agtcattggt | cgctgtagcc | gcaaaccccc | aggctcaagt | 1800 |
| | gatcctccg | cctcagcttc | ccagagtgt | gggattacag | gtgtgagcca | ctgcccttgc | 1860 |
| | ctggcacttt | taaaaaccac | tatgtaaggt | cagggtccagt | ggcttccaca | cctgtcatcc | 1920 |
| 5 | cagtagtttg | ggaagccgag | gcagaaggat | tgtctgaggc | caggagtttg | agaccagcat | 1980 |
| | gggtaacata | gggagacccc | atctctacaa | aaaatgcaaa | aagttatccg | ggcgtggggt | 2040 |
| | ccagcatctg | tagtcccagc | tgctcgggag | gctgagtggg | aggatcgctt | gagcccggga | 2100 |
| | ggtcatggct | gcagtgtgct | gtgattgtac | catcgactc | cagcctgggc | aacagagtga | 2160 |
| | gacctgtct | caaaaaaaaa | aaaaaaaaaa | gaaggagaag | gagaagagaa | gaagaaggaa | 2220 |
| 10 | gaaggaaaga | gaagaagaag | gaagaaggaa | gaaagaagga | gaaggaggcc | tgctaggtgc | 2280 |
| | taggttagact | gtcaaattct | agagcaaaat | gaaaataaca | aagttttaa | gggaaagaaa | 2340 |
| | aaccccagct | ctttggactt | ccttaggcct | gaacttcac | tcaagcagct | tccttccaca | 2400 |
| | gacaagcgtg | tatggagcga | gtgagttcaa | agcagaaagg | gaggagaagc | aggcaagggt | 2460 |
| | ggaggctgtg | ggtgacacca | gccaggaccc | ctgaaaggga | gtgggtgttt | tcctgcctca | 2520 |
| 15 | gccccacgct | cctgccggtc | ctgcacctgc | tgtaaccgtc | gatgttggtg | ccaggtggcc | 2580 |
| | acctgggaag | gatgctgtgc | agggggcttg | ccaaactttg | gtgggtttca | gaagccccag | 2640 |
| | gcacttgtgg | caggcacaa | tacagccct | ccccaaagat | gccccacgtc | ttctcctgga | 2700 |
| | acctgtgaat | gtgtcaccgc | caaggcagag | gctggtgaag | gctgcaggtg | gaatcacggc | 2760 |
| | tgccagtcag | ccgatcttaa | ggtcatcctg | gattatctgg | tgggcctgat | atggccacaa | 2820 |
| 20 | gggtccctag | aagtgtgaga | gggaggcagg | ggagagtcag | agaggggacg | tgagaaggac | 2880 |
| | cactggccac | tgctggcttt | gagatggagg | agggggtccc | cagccaagga | atgggggcag | 2940 |
| | ccgtccatg | ctggaaaagc | aagcaatcct | ccccggtcct | gagggcacac | ggcctgccc | 3000 |
| | acgcctcgat | ttcaggccag | tgggacctgt | ttcagctttc | cggcctccag | agctgtaaga | 3060 |
| | tgatgcgttt | gtgttcagcc | actaagctgc | agtgattcgt | cacagcagca | aatggaatag | 3120 |
| 25 | cagtacaggg | aaatgaatac | agggacagtt | ctcagagtga | ctctcagccc | accctggg | 3179 |

001260 972350